```
OM of: US-08-711-417C-165 to: PIR_71:*    out_format : pfs

Date: Aug 28, 2002 10:06 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_n2p.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/6228611/runat_28082002_100210_13574/app_query.fasta_1.1639
 -DB=PIR_71 -QFMT=fastan -SUFFIX=rpr -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.000
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1  -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=1 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=6228611_@CGN1_1-213 -NCPU=6
 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-08-711-417C-165
Query length: 1551
Database: PIR_71:*
Database sequences: 283138
Database length: 96089334
Search time (sec): 98.600000
```
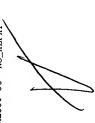
score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
pir2:A56355	+	2449.00	2899.19	4.1e-154	518	! DNA-binding protein Ikaros form
pir2:B56229	+	1938.50	2295.23	2.2e-120	427	! lymphoid transcription factor I
pir2:I59572	+	1933.50	2289.22	4.7e-120	431	! Ikaros DNA binding protein - mo
pir2:A56229	+	1607.00	1902.69	1.7e-98	392	! lymphoid transcription factor I
pir2:JE0288	+	323.00	374.61	1.3e-13	671	! krueppel-type zinc finger prote
pir2:S22954	+	321.50	374.53	1.6e-13	553	! finger protein zfp-37 - mouse
pir2:S51037	+	321.00	372.18	1.8e-13	675	! zinc-finger protein - African c
pir2:I48689	+	311.00	360.84	8.2e-13	636	! gene NK10 protein - mouse
pir2:S10245	+	309.00	362.31	1.1e-12	411	! finger protein, testis - mouse
pir2:S26823	+	308.00	355.23	1.3e-12	803	! zinc finger protein ZNF43 - hum
pir2:G02075	+	307.00	356.68	1.5e-12	595	! transcription repressor zinc fi
pir2:S35305	+	305.50	348.80	2.0e-12	1191	! zinc finger protein ZNF91 - hu
pir2:JH0500	+	305.00	354.03	2.0e-12	614	! zinc finger protein 29 - mouse
pir2:A38073	+	304.50	354.92	2.1e-12	519	! transcription activator ZNF35
pir2:A40751	+	304.00	354.92	2.3e-12	485	! finger protein MZF1 - human
pir2:B32891	+	303.50	351.74	2.6e-12	651	! finger protein 2, placental - h
pir2:A29253	+	303.50	350.40	2.6e-12	758	! finger protein hunchback - frui
pir2:A32891	+	301.00	352.46	3.6e-12	428	! finger protein 1, placental - h
pir2:A37107	+	298.50	346.82	5.4e-12	580	! spermatogenesis protein Zfp-35
pir2:I38600	+	298.00	348.10	5.7e-12	469	! zinc finger protein ZNF135 - hu
pir2:S00647	+	295.50	335.84	9.5e-12	1350	! finger protein - African clawe
pir2:S06556	+	294.50	344.53	9.6e-12	439	! finger protein (clone X1cOF28)
pir2:I39311	+	294.00	341.61	1.1e-11	572	! Kruppel-type zinc finger protei
pir2:I38599	+	291.00	342.42	1.6e-11	348	! zinc finger protein ZNF134 - hu
pir2:A48157	+	288.50	335.02	2.5e-11	576	! renal transcription factor Kid-
pir2:S05447	+	288.50	334.60	2.5e-11	604	! finger protein glass - fruit fl
pir2:T08674	+	287.50	335.15	2.8e-11	496	! probable finger protein DKFZp56
pir2:A34612	+	286.50	331.11	3.4e-11	686	! zinc finger protein ZNF7 - huma
pir2:T14757	+	286.00	330.33	3.7e-11	701	! hypothetical protein DKFZp572C1
pir2:I38598	+	285.50	331.26	3.9e-11	589	! zinc finger protein ZNF132 - hu
pir2:I37570	+	285.00	329.24	4.3e-11	693	! zinc finger protein - human (fr
pir2:I50620	+	284.50	324.02	5.0e-11	1173	! procKr2- chicken (fragment)
pir2:JC5146	+	279.00	323.48	1.1e-10	594	! arylphorin gene-specific bindin
pir2:A57785	+	278.50	322.04	1.2e-10	654	! finger protein ZNF133 - human
pir2:S08686	+	278.00	322.47	1.2e-10	582	! finger protein ZFP-36 - human
pir2:S05548	+	277.50	318.91	1.4e-10	816	! gap protein hunchback - fruit f
pir2:S03677	+	276.50	323.39	1.5e-10	428	! finger protein (clone mkr3) -
pir2:S68858	+	276.00	318.39	1.7e-10	707	! finger protein - mouse
pir2:B34612	+	275.50	320.11	1.8e-10	543	! zinc finger protein ZNF8 - huma
pir2:JH0501	+	274.50	319.50	2.1e-10	509	! zinc finger protein CKr1 - chid

```
pir2:S47071    +    273.00   319.86   2.5e-10   399  ! finger protein HZF3, Krueppe
pir2:S60520    +    272.00   320.16   2.9e-10   337  ! finger protein ZNF81.1 - hum
pir2:S46593    +    271.00   310.20   3.8e-10   914  ! finger protein AZF1, asparag
pir2:S47073    +    269.50   310.37   4.6e-10   732  ! finger protein HZF2, Krueppe
pir2:B39240    +    268.00   313.75   5.4e-10   407  ! finger protein mfg2 - mouse
```

seq_name: pir2:A56355

```
seq_documentation_block:
DNA-binding protein Ikaros form 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 01-Dec-2000
C;Accession: A56355
R;Molnar, A.; Georgopoulos, K.
Mol. Cell. Biol. 14, 8292-8303, 1994
A;Title: The Ikaros gene encodes a family of functionally diverse zinc finger DNA-bin
A;Reference number: A56355; MUID:95059058
A;Accession: A56355
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-518 <MOL>
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C;Keywords: alternative splicing; DNA binding
```

```
alignment_scores:
    Quality: 2449.00          Length:        521
      Ratio: 4.928              Gaps:          6
Percent Similarity: 95.393    Percent Identity: 89.827
```

alignment_block:
US-08-711-417C-165 x A56355

Align seg 1/1   to: A56355   from: 1   to: 518

```
    1 ATGGGATCTGACGAGGTCAAGACACATGTCTTTCCATCAGGGAAGGAAAG  50
      ||||||||:::||||||||||||||||||||||||||||||  ||||||||
    1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe  17

   51 CCCCCCTGTAAGGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC  100
      ||||||||||||||||||||||||||||||||||||||||||:::||||
   17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP  34

  101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAGCAAAGCTCCAAGAGTGAC  150
      |||||||||||||||||||||||||||||||||||||||:::||||||||
   34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp  50

  151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGATGATGAAGAGAA  200
      ||| :::|||||||||||||||||||||||||||||||||||||||||||
   51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerSerGluGluAs  67

  201 TGGGGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA  250
      ||||||||||||||||||||||||||||||||||||||||||||||
   67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM  84

  251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCCACAGGGACCAAGGC  300
      |||||||||||||||||||||||||||||||||||||||||||||||||
   84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly  100

  301 AGCTCGGCTTTGTCGGGAGTTGGAGCATTCGACTTCCTAACGGAAAACT  350
      |||||||||||||||||||||||||||||||||||||||||||||||
  101 SerSerAlaLeuSerGlyValGlyValGlyIleArgLeuProAsnGlyLysLe  117

  351 AAAGTGTGATATCTGTGGATCATTTGCATCGGCCCCAATGTGCTCATGG  400
      |||||||||||:::||||||||||||||||||||||||||||||
  117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV  134

  401 TTCACAAAAGAAGCCACACTGAGAACGGCCCTTCCAGTGCAATCAGTGC  450
      |||||||||||||||||||||||||||||||||  ||||||||||||||
  134 alHisLysArgSerHisThrGlyGluArgProPheCysAsnGlnCys  150

  451 GGGGCCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCA  500
      ||||||||||||||||||||||||||||||||||||||||||||||
```

```
1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.........G 1435
     ||||||||||||||||||||||||||||||||||
 464 .gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     |||||||||||||||||||||||||||||||||||||||||
 481 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 497

1486 AGCCAGGACCGTACGAGTTCTCGTCGCACATAACGCGAGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||
 498 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 514

1536 CTTCCACATGAGC 1548
     |::::||:::|||
 514 gTyrHisLeuSer 518

seq_name: pir2:B56229
```

```
alignment_scores:                           Length:        518
        Quality: 1938.50                    Gaps:            7
        Ratio:     4.728
Percent Similarity:  79.151     Percent Identity: 73.745

alignment_block:
US-08-711-417C-165 x B56229

Align seg 1/1  to: B56229  from: 1   to: 427

   1 ATGGATGCTGACGAGGTCAAGACATGTCTTCTCATCAGGAAGGAAAG 50
     ||||||||||||||||||||||||||||||||| |||||||||||
   1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluE 17

  51 CCCCCCTGAAGGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
     ||||||||||||||||||||||||||||||||||||||||:::|
  17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

 101 CCGAGGACCTCTCCACCACTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
     ||||||||||||||||||:::|||||||||||||||||||||
  34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

 151 AGAGTCGTGGCCAGTAAGTTAAAGTAGAGACTTCAGAGTGATGAAGAGAA 200
     ||| |||                              ........
  51 ArgGlyMet.......................... 53

 201 TGGGCGTGCCTGTGAAATGAATGGGAAGAATGTGCGGAGGATTTACGAA 250
     ............ 53

 251 TGCTTGATGCTCGGGAGGAGAAAAATGAATGGCTCCCACAGGACCAAGGC 300
     ............ 53

 301 AGCTCGGCTTTGTCGGAGTTGGAGGCATTGCAGTTCCTAACGGAAAACT 350
     ............ 53
```

```
 151 GlyAlaSerPheThrGlnLysGlyAsnLeuArgHisIleLysLeuHi 167
     |||||||||||||||||||||||||||||||||||
 501 TTCCGGGGAGGAAGCCCTTCAAATGCCACCTGTGCAACTACGCCTGCCGCC 550
 167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

 551 GGAGGGACGCCCTCACTGCCACCTGAGGACGGCACTCCGTTGGTAAACCT 600
     ||||||||||||||||||||||||||||||||||||||||
 184 rgArgAspAlaLeuThrGlyHisHisLeuArgThrHisSerValGlyLysPro 200

 601 CACAAATGGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
     ||||||||||||:::|||||||
 200 HisLysCysCysSerGlyTyrTyrLysSerGlyTyrLysArgSerSerLeuGl 217

 201 HisLysCysCysSerGlyTyrTyrLysSerGlyTyrLysArgSerSerLeuGl 217
 651 GGAACATAAAGAGACGCTGCCACAACTACTTGAAAGCATGGGCCTTCCGG 700
 217 uGluHisLysGlyLeuArgCysHisIleAsnTyrLeuGluSerMetGlyLeuProG 234

 701 GCACACTGTACCCAGTCATTAAAGAAGAAAACTAAGCACAGTGAAATGCA 750
     ||  :::||||||||||||||||||||:::||:::||||||||
 234 ly...MetTyrProValIleLysGluGluThrAsnHisAsnGluMetAla 249

 751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
     |||||||:::||||||:::|||||||||||||||||||||
 250 GluAspLeuCysLysIleGlySerGluGlyAlaGluArgSerleuValLeuAspArgLe 266

 801 AGCAAGTAATGTCGGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
     |||||||||||||||||||||||||||||||||||||||
 266 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 283

 851 GGGACAAGGGCCTGTCGGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
     ||||||        |||||||||||||:::||||||||||||
 283 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 299

 901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
     |||      :::|||||||||||||||||||||||||||||
 300 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 315

 951 CATCAACTACCTGGGGCGGAGTCCTGGCGCCCCTGGTCCAGACGCCCCC 1000
     |||||||||||||||||||||||||||||||||||||||
 315 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 332

1001 CGGGGCCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
     ||||:::||||||||||||||||||  ||||||||||||||
 332 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 348

1051 AGG...CGCTCGGAGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
     |||     ||:::|||    |||||||||||||||||||||
 349 LysProProSerGlyProProArgSerAsnHisSerAlaGlnAsp... 364

1098 CGCCGTGGAGTACCTGCTGCTGCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
     ||:::||||||||||||||  |||   ||||||||||||
 365 .AlaValAspAsnLeuLeuLeuLeuSerLysLysSerValSerSerG 381

1148 AGGCGAGGGGGTCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     |||||||||||||:::|||||||||||||||||||||||
 381 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 397

1198 AGCAACAACGAGGAGCAGCCAGCCGGTCTTATCCTACCTGACCAACCACAT 1247
     |||||||||||||||||||||||||||||||||||||
 398 SerAsnAsnAlaGluGluGlnProAlaGlyLeuIleLeuThrAsnHisIl 414

1248 CGCCCGACGCGGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |:::||||::::   :::|||||||||:::|||||||
 414 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluHisArgAlaT 431

1295 ACGACCTGCTCGCGCCGCCGCCTCCGAGAACTCGCAGGACGCGGCTCCGCGTG 1344
     ||:::::|||:::|||||||||||||||||||||||||
 431 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 447

1345 GTCAGCACCAGCCGGGGAGCAGATGAGAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||||||:::||||||||||||||||
 448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464
```

```
  53 ..............................................  53

 351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400

  53 ..............................................  53

 401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
  54 ..............GlyGluArgProPheGlnCysAsnGlnCys       63

 451 GGGGCCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGTGCA    500
  64 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi   80

 501 TTCCGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGGCTGCGCC    550
  80 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA   97

 551 GGAGGGACCCCCTCACTGGCCACCTGAGGACGCCACTCCGTTGGTAAACCT 600
  97 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro  113

 601 CACAAATGGATATTGTGGCCCAAGCTATAAACACGCGAACGTCTTTAGA   650
 114 HisLysCysGlyTyrCysGlyLysTyrLysGlnArgSerSerLeuGl     130

 651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG  700
 130 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG  147

 701 GCACACTGTACCCAGTCATTAAAGAGAAACTAAGCAGTGAAATGGCA     750
 147 ly...MetTyrProValIleLysGluGluThrAsnHisAsnGluMetAla  162

 751 GAAGACCTGTCCAAGATAGGATCAGAGATCTCTGCTGTGGACAGACT     800
 163 GluAspLeuCysLysIleGlyAlaGluSerLeuValLeuAspArgLe     179

 801 AGCCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
 179 uAlaSerAsnValAlaLysArgArgLysSerSerMetProGlnLysPheLeu. 195

 851 GGGACAAGGGCCTGTCCGACACGCCCACGACAGTGCCACGTACGAGAAG   900
 196 ..AspLysCysLeuSerAspAspMetProTyrAspSerAlaAsnTyrGluLys 211

 901 GAGAACGAAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
 212 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl  227

 951 CATCAACTACCTGGGGCCAGTCCTGCCGCCGCTGGTGCAGACGCCCC    1000
 227 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP  244

1001 CGGCGGTTCCGAGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC   1050
 244 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis  260

1051 AGG...CGCTCGGAGGGCCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
 261 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp..  276

1098 CGCCCTGGAGTACCTGCTGTCTCCAAGGCCAAGTTGGTCCCTCGG      1147
 277 .AlaValAspAsnLeuLeuLeuSerLysLysSerValSerSerG        293

1148 AGCCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG  1197
 293 luArgGluAlaSerProSerArgSerCysGlnAspSerThrAspThrGlu  309

1198 AGCAACAACGAGGAGCAGCCGCAGCCGGTCTTATCTACCTGACCAACCACAT 1247
 310 SerAsnAlaGluGluGlnProAlaSerGlyLeuIleTyrLeuThrAsnHisII 326
```

```
1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
 326 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGlnLeuArgAlaT  343

1295 ACGACCTGCTGCGCGCCGCTCCCAGACTCGGCAGGACGCGCTTCCGGCTG 1344
 343 YrGluValLeuArgAlaAlaSerGluAsnSerGlnAsnProAlaPheArgVal 359

1345 GTCAGCACCAGCGGAGGCAGGAGAGATGAAGGTGTACAAGTGCCAACACTGCCG 1394
 360 ValSerThrSerGlyGlyArgArgGlnLnLeuLysValTyrLysCysGluHisCysAr 376

1395 GGTGCTCTTCCTGGATCACGTCATGTACCAACATCCACATGGCGTGCCACG 1444
 376 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHissG 393

1445 GCTTCCGTGATCCTTTGAGTGCAACATGTGCGCTACCAGCCAGGAC     1494
 393 lyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp  409

1495 CGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCCACAT 1544
 410 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArgTyrHisLe  426

1545 GAGC 1548
     :|||
 426 uSer 427
```

seq_name: pir2:I59572

seq_documentation_block:
Ikaros DNA binding protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 01-Dec-2000
C;Accession: I59572
R;Georgopoulos, K.; Moore, D.D.; Derfler, B.
Science 258, 808-812, 1992
A;Title: Ikaros, an early lymphoid-specific transcription factor and a putative media
A;Reference number: I59572; MUID:93068267
A;Accession: I59572
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-431 <RES>
A;Cross-references: GB:L03547; NID:g198286; PIDN:AAA66193.1; PID:g198287
C;Genetics:
A;Gene: Ikaros
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

alignment_scores:
          Quality: 1933.50   Length: 521
            Ratio: 4.716   Gaps: 7
Percent Similarity: 78.695   Percent Identity: 73.321

alignment_block:
US-08-711-417C-165 x I59572

Align seg 1/1 to: I59572   from: 1 to: 431

```
  1 ATGGATGCTGACGAGGGGTCAAGACCATGTCTTCTCATCAGGGAAGGAAAG  50
    |||||::::||||||||||||||||||||     ||||||||||||
  1 MetAspValAspGluGlnAspMetSerGlnInValSerGlyLysGluSe   17

 51 CCCCCTGTAAGCGATACTCCAGATGAGGCGGCCATGCCGATCC        100
    |||||||||::|||||||||||||||||||||||::::|
 17 rProProValSerAspThrProAspGluGlyIlyAspGluProMetProValP 34

101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAGAAAGCTCCAAGAGTGAC 150
    |||||||||||||||||||||||||||||||||||
 34 roGluAspLeuSerThrThrSerGlyAlaGlnInAsnSerLysSerAsp   50

151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAAGAGACTCAGAGTGATGAAGAA 200
```

```
 262 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp..       277
     |||                :::
1098 CGCCGTGGAGTACCTGCTGCTGCTCCAAGGCCAAGTTGGTGCCCTCGG          1147
     ||||||:::        ||||||||||||||||           ||
 278 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG        294

1148 AGCGGAGGCGGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG        1197
     |||||||||||||||||||||||||||||||||||||||||||||||
 294 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu        310

1198 AGCAACAACGAGGAGCAGCCGAGCGGTCTTATCTACCTGACCAACCACAT        1247
     |||||      ||||||||||||||||||||||||||||||||||||
 311 SerAsnAlaGluGluGlnGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl      327

1248 CGCCGACGCGCGCCAACGC...GTGCGCTCAAGAGGAGCACCGCGCCT        1294
     ::|||:::         :::|||:|||||:|||||||||||||
 327 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGlnArgAlaT          344

1295 ACGACCTGCTGCGCGGCCTTCCGAGAACTCGCAGAACGCGGCTCCGCGTG        1344
     ||:::||:||||||||||||||||||||:||||||||||||::||||||
 344 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAsnAlaPheArgVal        360

1345 GTCAGCACCAGCCGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG        1394
     ||||:||||||||||||||||||||||||||||||||||||||||||||
 361 ValSerThrSerGlyGluGlnMetLysValTyrLysCysGluHisCysAr         377

1395 GGTGCTCTTCCTGATCACGTCATGTACACCATCCACATG.........G      1435
     |||||||||||||||||||||||||||||||||||||||         |
 377 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG        394

1436 GCTGCCACGGCTTCCGTGATCCTTTGAGTGCAACATGTGCGGCTACCAC        1485
     |||||||||||||||||||||||||||||||||||||||||||||||
 394 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis        410

1486 AGCCAGGACCCGGTACGAGTTCTCGTCGCCACATAACCGCGAGGGAGCACCG       1535
     |||||||||||||||||||||||||||||||||||||||||||||||
 411 SerGlnAspProValPheSerSerHisIleThrArgGlyGluHisAr          427

1536 CTTCCACATGAGC        1548
     |:::||||:::|||
 427 gTyrHisLeuSer        431
```

seq_name: pir2:A56229

```
      |||           :::
  51 ArgGlyMet.......................................       53

 201 TGGGCCGTGCCTGTGAAATGATGGGGAAGAAATGTGCGGAGGATTTACGAA       250

  53 ...........................................         53

 251 TGCTTGATGCCTCGGAGAGAAAAATGAATGGCTCCCACAGGACCAAGGC        300

  53 ...........................................         53

 301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT        350

  53 ...........................................         53

 351 AAAGTGTGATATCTGTGGGATCATTGCATCTGGGCCCAATCGTCTCATGG       400

  53 ...........................................         53

 401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCCTTCCAGTCGAATCAGTGC       450
                    ||||||||||||||||||||||||||
  54 ...........GlyGluArgProPheGlnCysAsnGlnCys           63

 451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTTCCGGCACATCAAGCTGCA      500
     ||||||||||||||||||||||||||||||||||||||||||||||||
  64 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi      80

 501 TTCCCGGGAGAAGCCCTTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC      550
     |||||||||||||||||||||||||||||||||||||||||||||||
  80 sSerGlyGluLysProPheLysCysSerHisLeuCysAsnTyrAlaCysArgA     97

 551 GGAGGGACGCCCCTCACTGGCCACCTGAGGACGGCCACTCCGTTGGTAAACCT     600
     |||||||||||||||||||||||||||||||||||||||||||||||
  97 rgArgAspAlaLeuThrGlyHisLeuArgThrAlaThrProLeuValGlyLysPro    113

 601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA      650
     ||||||||||||||||||||||||||||||||:::||||||||||
 114 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerLeuGl         130

 651 GGAACATAAAGAGCCGCTGCCACACTACTTGGAAAGCATGGGCCTTCCGG      700
     |||||||||||||||||||||||||||||||||||||||||||||||
 130 uGluHisLysLysGluArgCysHisSerAsnTyrLeuGluSerMetGlyLeuProG    147

 701 GCACACTGTACCCAGTCATTAAAGAAGAAAACTAGCACAGTGAAATGGCA       750
                 |||||||||::::||||||||||||||
 147 ly..ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla       162

 751 GAAGACCTGTCGAAGATAGGATCAGAGAGATCTCTGCTGGACAGACT        800
     ||||||||||||||||||||||||||::::||||||||||||||
 163 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe       179

 801 AGCAAGTAATGTCGCCAAACCTAAGAGCCTCTATGCCTCAGAAATTTCCTTG     850
     ||||||||||||||||||||||||||||||||||||||||||||||
 179 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG       196

 851 GGACAACAAGGGCCTGTCCGACACGCCCCTACGACAGTGCCACTACCGAGAG     900
     ||||||||||||||||||||||||::::||||||||||||||||||
 196 lyAspLysCysLeuSerAspProTyrAspSerAlaAsnTyrGluLys         212

 901 GAGAACGAAATGACAAGTGAAGTCCCCACGTGATGACCAAGCCATCAACAACGC    950
     ||||||||    :::||||   ||||||||||||||||||||||||
 213 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl      228

 951 CATCAACTACCTGGGGGCCGAGTCCCTGCCCCGCCGCTGGTGCCAGACGCCCC    1000
     ||||||||||||||||||||||||||||||||||||||||||||||
 228 aIleAsnTyrLeuGlyAlaGluProLeuProValGlnThrProP           245

1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC     1050
     ||||||||||||||||||||||||||||:::||||||||||||
 245 roGlySerGluValValProValIleSerProMetTyrGlnLeuHis        261

1051 AGG...CGCTCGGAGGGGCACCCCCGCGCTCCAACCACTCGGCCCAGGACAG    1097
     :::              ||||||||||||||||||||||||
```

```
                                                        ||||||
                                          161 ..........AspLysCysLeuSerSerAspMetProTyrAspSerAlaAs 173

  1 ATGGATGCTGACGAGGGTCAAGACATGTCTTCTCATCAGGGAAGGAAAG  50    891 GTACGAGAAGGAGAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCA 940
    ||||||::::||||||||||||||||||||||||                         |||||||||| ::::|||||| |||||||||||||||
  1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe  17    173 nTyrGluLysGlu...AspMetMetThrSerHisValMetAspGlnAlaI 189

 51 CCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC  100    941 TCAACAAGCCATCAACTACCTGGGGGCCGAGTCCCTGCGCCCCGCTGGTG 990
    |||||||||||||||||||||||||||||||||:|                        ||||||||||||||||||:::|||||||||||
 17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP  34    189 leAsnAsnAlaIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuVal 205

101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCCAAAGTTCCAAGAGTGAC 150    991 CAGACGCGCCCCGGCGGGTTCCGAGGTGGTGTCCGGTCATCAGCCCGATGTA 1040
    |||||||||||||||:::||||||||                               |||||||||||||||:||::||||||||||||||
 34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp  50    206 GlnThrProProGlySerSerGluValValProValIleSerSerMetTy 222

151 AGA...GTCGTGGCC..................                   162    1041 CCAGCTGCCACAGG...CGCTCGGAGGGCACCCCGGCTCCAACCACTCGG 1087
    |||  ::||||||                                            ||||||||:::  ||||:||||||||||||
 51 ArgGlyMetValAlaTyrGlyAlaAspGlyPheArgAspPheHisAlaIl  67    222 rGlnLeuHisLysProProSerAspGlyProArgSerAsnHisSerA 239

163 ............AGTAAATGTTAAAGTAGAGACTCAGAGTG            190    1088 CCCAGGACAGCCCGTGGAGTACCTGCTCTCCAAGGCCAAGTTG 1137
    ||||||||||||||||||                                       |||||| ||||:: :||||||||||||||
 67 eIleSerAspArgGlyMetProSerAsnValLysValGluThrGlnSerA  84    239 laGlnAsp...AlaValAspAsnLeuLeuSerLysAlaLysSer 254

191 ATGAAGAGAATGGGCGTGCCTGTGAAATGAATGGGAAGAATGTGCCGAG   240    1138 GTGCCCTCGGAGGCGGAGGCGTCCCCAGCAACAGCTGCCAAGACTCCAC 1187
    ||||||||||||||||||||||||||                              ||| |||||||||||||||||||||||||
 84 spGluGluAsnGlyArgAlaCysGluMetAsnGlyLysGluCysAlaGlu 100    255 ValSerSerGluArgGluAlaSerProSerAsnSerCysGlnAspSerTh 271

241 GATTTACGAATGCTTGATGCCTCGGAGGAGAAAATGAATGGCTCCACAG   290    1188 GGACACCGAGAGCAACAACGAGGAGGAGCGCAGCCGGTCTTATCTACCTGA 1237
    |||||||||||||||||||||||||||                             ||||||||||||||||||||||||||
101 AspLeuArgMetLeuAspAlaSerGluGlyGluLysMetAsnGlySerHisAr 117    271 rAspThrGluSerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuT 288

291 GGACCAAGGCAGCTCGGCTTGTCGGGAGTTGGAGGCATTCGACTTCCTA   340    1238 CCAACCACATCGCCCGACGCGGCGCAACGC...GTGTCGCTCAAGGAGGAG 1284
    ||||||||||||||||||||||||||                              ||||||| :::||||||||  ::::::||||||||||||
117 gAspProGlySerSerAlaLeuSerGlyValGlyIleArgLeuProA    134    288 hrAsnHisIleAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGlu 304

341 ACGGAAAACTAAAGTGTGATATCTGTGGATCATTTGCATGGGCCCAAT    390    1285 CACCGGCGCCTACGACCTGCTCGCGCCGCCCTCCGAGAACTCGCAGGACGC 1334
    ||||||||||||:::||||||||                                  ::||||||||||:::|||||||||||||||||||||
134 snGlyLysLeuLysCysAspIleCysGlyIleIleGlyIleClyProAsn 150    305 GlnArgAlaTyrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAl 321

391 GTGCTCAGGTTCACAAAAGAGACCACCAGGAACGGCCCTTCCAGTG     440    1335 GCTCCCGCTGGTCAGCACCACCAGCGGGAGCAGATGAAGGTGTACAAGTGCG 1384
    ||||||||||||||||||                                       |::|||||||::|||||||||||||||||||||||||
151 ValLeuMetValHisLysArgSerHisThr..............        160    321 aPheArgValValSerThrSerGlyGluGlnLeuLysValTyrLysCysG 338

441 CAATCAGTGCGGGGCCTCATTCACCCAGAGAGGGCAACCTGCTCCGGCACA 490    1385 AACACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACCACCATCCACATG 1434
160 .......................                                  ||||||||||||||||||||||||||||||
                                                         338 luHisCysArgValLeuPheLeuAspHisValMetTyrThrIleHisMet 354

491 TCAAGCTGCATTCCGGGGAGAGCCCTTCAAATGCCACCTCTGCAACTAC   540    1435 GGCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTCGCGCTACCA 1484
160 .......................                                  |||||||||||||||||||||||||||||||
                                                         355 GlyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHi 371

541 GCCTGCCGCCGAGGGACGACGCCCCTCACTGGCCACCTGGACGGCCACTCCGT 590    1485 CAGGCAGGACCGGTACGAGTTCTCGTCGCACATAACCGCGAGGGGAGCACC 1534
160 .......................                                  |||||||||||||||||||||||||||||||
                                                         371 sSerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisA 388

591 TGGTAAACCTCACAAATGTGGATAATTGTGGCCGAAGCTATAAACAGCGAA 640    1535 GCTTCCACATGAGC 1548
160 .......................                                  ||::|||::::|||
                                                         388 rgTyrHisLeuSer 392

641 CGTCTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATG 690
160 .......................

691 GGCCCTTCCGGGCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAG  740
160 .......................

741 TGAAATGGCAGAAGACCTGTGCAAGATAGGATCAGAGAGGATCTCCGTGC  790
160 .......................

791 TGGACAGACTAGCAAGTAATGTCGCCAAACGTAAGAGCTTCTATGCCTCAG 840
160 .......................

841 AAATTTCTTGGGGACAAGGGCCTGCCGACACGCCCTACGCCGACAGTGCCAC 890
```

seq_name: pir2:JE0288

Wed Aug 28 10:05:20 2002

us-08-711-417c-165.rpr

```
A:Residues: 1-671 <KAT>
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

alignment_scores:
       Quality:   323.00          Length:       550
         Ratio:    1.166            Gaps:        21
Percent Similarity:  50.364   Percent Identity:  22.364

alignment_block:
US-08-711-417C-165 x JE0288  .

Align seg 1/1  to: JE0288  from: 1  to: 671
```

```
103  GAGGACCTCTCCACCCCTCGGGAGGACAGCAAAGCTCCAAG..AGTGA  149        700  GGCACACTGTACCCAGTCATTAAAGAAGAAAACTAAGCACAGTGAAATGGC  749
     ::::   ::  ||:||  ||:::  ||::  ||:::| |||:|  ::|||         ||:|:||||         |||||||||
73   GluSerLysAspGlyThrGlnCysGlyGluThrSerGlnIleGlnAsn  89         322  GlySerLeuGln........ArgHisGluThrThrHisSerAlaGluLy  335

150  CAGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGA  199        750  AGAAGACCTGCAAGATAGGATCACAGAGATCTCCGTGCTGACAGAC  799
     : ::||||||:||||::  ||        :||       |||:||:::          ||||||||:         |||         |||
89   pSerIleValThrLeuLysAsnThrLeuProGlyValGlyProCysGluSers  106       335  sProTyrAlaCysGlnCysGlnLysGlyLysAlaPheHisHisLeuGlySerP  352

200  ATGGGCTGCCTGTGAAATGAATGGGAAGAATGTGCGGAGGATTTACGA  249        800  TAGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCCTCAGAAATTT...  846
     |:  || |||  :::|||:|  |::||||::||:|||||| |:||  ::          ::   ::::::  |:::  :||  ||:::|||
106  erMetArgGlyLeuLysValMetGlyValHisSerLeuAsnCysTyrIle  122       352  heGlnArgHisMetIleArgHisThrGlyAsnGlyProHisLysCysLys  368

250  ATGCTTGATGCCTCG........................GGAGAGAA  272       847  CTTGGGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGA  896
     ::  |||  ::                            |||||||         ||| ||| |||  |::||:|  :::||:|||:|
123  ArgValGlyAlaGlyHisLysLysProHisGluTyrHisGluCysGlyGluLy  139       369  IleCysGlyLysGlyPhe...AspCysProSerSerLeuGlnSerHisGl  384

273  AATGAATGCCTCCCACAGGACCAAGGCAGCTCG................  306       897  GAAGGAGAACGAAAATGATGAGAGTCCCACGTGATGACCAAGCCATCAACA  946
     : :||||||||::|||||||||||||:|||                          | :::         |||         ::::::::
139  sProAsp...ThrHisLysGlnArgGlyThrLysAlaPheSerTyrHisAsnS  155       384  uArgThrHisThrGlyGluLysProTyrGluCysLysGlnCys...GlyL  400

306  ....................................................  306       947  ACGGCCATCAACTACCTGGGGCCGGCCGAGTCCCTGCCCCCGCTGGTCAGACG  996
                                                                      :|||||:|||::  ::::::  |:::  |||:||||
155  erPheGlnThrHisGluArgLeuHisThrGlyLysLysProTyrAspCys  171       400  ysAlaLeuSerHisArgSerPheArgSerHisMetIleMetHisThr  416

307  ............GCTTTGTCGGGAGTTGGA.................  324       997  CCCCCGGCGGTTCCAG...........................GTGGTCCC  1022
                |||:|:|||:|||||||||                                 ||| |||
172  LysGluCysGlyLysSerPheSerSerLeuGlyLysAsnLeuGlnLeuArgHisMe  188       417  GlyAspGlyProHisLysCysLysValCysGlyLysAlaPheValTyrPr  433

325  .GGCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTGTGGGATCA  373       1023 GGTCATCAGCCCGATGTACCAGTCGCACAGGCGCTCGGAGGGCACCCCGC  1072
     :|||||:|::  |:||::|||||||  ||||||:|:||:||||:|||           :::|:||   |||:|:|||:|
188  tAlaValGlnArgGlyAspGlyProTyrLysCysLysLeuCysGlyLysLysA  205       433  oSerVal..........PheGlnIleArgHisGluArgThr.......  442

374  TTTGCATCGGGCCCAATGTCCATGGTTCACAAAGAGCCACACACTGGA  423       1073 GCTCCAACCACTCGGCCCAGGACACGCGCCGTGGAGTACCTGCTGTGCTC  1122
     ||::::||  ||:::||  ||::|:|||||:|||||:::||||||||||                :::|||:||||||||
205  laPheTrpProSerLeuHisMetHisGlyArgThrHisThrGly  221       443  ........HisThrAlaGlu........IleProTyr.........  449

424  GAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCCATTCACCCAGAAGGG  473       1123 TCCAAGGCCAAGTTGGTGCCCTCGGAGCGGCGAGGCGCTCCCCGAGCAACAG  1172
     |||:::||:||::::|||||:||  ::|||:::  ::                 ||:::|||    ::: |||  |||  |||
172  GluLysProTyrGluCysGlnCysGlnLeuCysGlnCysGlyGlnLysLysA  238       450  ...LysCysLysGlnCysGlyLysLysAlaTyrArgIleSerSerLeuArg  465

424  GCCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTGTGGGATCA  238       1173 CTGCCAAGACTCCACGGACACCGAGACCAACCAGGAGGACCAGCAGCCAGCG  1222
     |||::::|||  :::||::|||||||  ||||:|:||:|||||:|||          :::|:|||:|||    |||         |||
222  GluLysProTyrGluCysGlnCysGlnLeuCysSerLysLysArgSerPheTyrSe  238       465  gArgHisGluThrThrHisThrGlyGluLysProTyrLysCysThrLeuG  482

474  CAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAAGCCCTTCAAAT  523       1223 GTCTTATCTACCTG.............................ACCAACCAC  1245
     :::  ||||:::|||  ::        ||::|||||||||||||:|                                      ::
238  rSerTyrLeuArgHisGluArgThrHisThrGlyValHisGluLysProTyrGluC  255       482  lyLysAlaPheIleAspPheCysSerPheGlnAsnHisLysThrHis  498

524  GCCAACCTGCCAACTACCGCCTGCCGCCGAGGACGGCCCTCACTGGCCAC  573       1246 ATCGCCCGAGCGGCGCAACGCGTGTCCTCAAGGAGGAGCACCGGCGCCTA  1295
     ||         |||  |||::::|||||||||::||||               :::|||    ||||||         :::||:
524  GCCAACCTGCCAACTACCGCCTGCCGCCGAGGAC   271       499  ThrGlyGluLysPro........TyrGluCysGluCysGlyLysAlaPh  513

255  YsLysGlnCysSerLysLysArgSerPheAlaPheProPheTyrSerSer  271       1296 CGACCTGCTGCGCGGCCGCCTCCGAGAACTCGCAGGACGCGCCTCCGGTGG  1345
                                                                     :::|||:|||    ||||||||||||||||
                                                                513  eSerArgPheArgTyrLeuSerArgHisLysArg................  524

574  CTGAGGACGGCCACTCCGTTGGTGAAACCTCACAAATGTGGATAATTGTGGCCCG  623       1346 TCAGCACCAGCGGGAGCCAGATGAAGGTGTACAAGTGCGAACACTGCCGG  1395
     |||||||||||||         ::: ||||||::|||||:::               :::|:||||    |||:|:||::|| ||||||
238  rSerTyrLeuArgHisGluArgThrHisThrGlyValHisGluLysProTyrGluC  255       499  TyrGluCysGluCysGlyLysAlaPh  513

524  GCCAACCTGCCAACTACCGCCTGCCGCCGAGGACGGCCCTCACTGGCCAC  573       1346 TCAGCACCAGCGGGAGCCAGATGAAGGTGTACAAGTGCGAACACTGCCGG  1395
                                                                525  ..ThrHisThrGlyGlu......LysProTyrGluCysLysThrCysArg  538

272  GluArgThrHisThrGlyThrGlyThrGlyProTyrLysCysLysCysGlnCysSerLy  288       1396 GTGCTCTTCCTGGATCACGTCATGTACACCATCCACATGGCTGCCACGG  1445
                                                                    |||         :::|||   ::::::
                                                                539  LysAlaPheGlyHisTyrAspAsnLeuLysValHisGluArgIleHisSe  555

624  AAGCTAT................................AAACAGCGAACGTCTTTAG  649       1446 CTTCCGTGATCCTTTGAGTGCAACATGTCGCGCTACCACAGCCAGGACC  1495
     ::  ||||||||                                        :  ||:|||||||:  |||    |||    :::||
555  rGlyGluLysProTyrGluCysLysGluCysGly..........LysA  568       539  LysAlaPheGlyHisTyrAspAsnLeuLysValHisGluArgIleHisSe  555

288  sAlaPheProAspSerSerCysSerLeuIleHisGluArgThrHisThrG  305       1496 GGTAGAGTTCTGTCGCCACATAACCGCGAGGGGAGCACCGCCTTCCACATG  1545
                                                                     :::|||:|:|||    |||:|||||||||||
650  AGGAACATAAAGACGCCTGCCACAACTACTTGGAAAGACATGGGGCCTTCCG  699       1496 GGTAGAGTTCTGTCGCCACATAACCGCGAGGGGAGCACCGCCTTCCACATG  1545
     |||  ::    :::||||||||||||||                         568  laPheSerTrpLeuThrCysPheLeuArgHisGlu...ArgIleHisMet  583

305  lyGluLysProTyrThrCysLysGlnCysGlyLysLysAlaPheSerValSer  321
```

```
seq_name: pir2:S22954

seq_documentation_block:
finger protein zfp-37 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 01-Dec-2000
C;Accession: S22954
R;Burke, P.S.; Wolgemuth, D.J.
Nucleic Acids Res. 20, 2827-2834, 1992
A;Title: Zfp-37, a new murine zinc finger encoding gene, is expressed in a developmental
A;Reference number: S22954; MUID:92310982
A;Accession: S22954
A;Molecule type: mRNA
A;Residues: 1-553 <BUR>
A;Cross-references: EMBL:X64413; NID:g55474; PIDN:CAA45758.1; PID:g55475
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C;Keywords: DNA binding

alignment_scores:
     Quality:  321.50          Length:   532
     Ratio:    1.204           Gaps:      27
Percent Similarity:  50.188    Percent Identity:  24.624

alignment_block:
US-08-711-417C-165 x S22954

Align seg 1/1 to: S22954 from: 1 to: 553
```

```
  28 TCTTTCTCATCAGGGAGGAAGAAAGCCCCCCTGTAAGCGATACT........   69
     |||  ||| |||||||::|||          ||||||||
 105 SerSerSerArgGlyLysAsnSerAsnGlnAsnSerAspSerLeuLysLy   121

  70 ....CCAGATGAGGCGATGAG.............................   87
         ||||| |||||||:|||
 121 sLysProAspThrAlaAsnGluHisArgLysSerLeuSerHisSerAlaS   138

  88 ........CCCATGCGATCCCGAGGACCTCCACCACC.............  120
              ::|||          ||  :||
 138 erAspValAsnLysAspGluIleProThrArgLysLysCysAspLysLeu   154

 121 TCGGGAGGACAGCAAAGTCCAAGAGTGACGAGTCGTGGCCAGTAATGT..   170
     |||||||||| |||:|||:|||||:|
 155 ProAsnAsnLysLeuSerAspLysGlyAspLys.................   165

 171 TAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCCGTGCCTGTGAAATGA   220
     :::|||||||::::          ::|||  :|
 166 .........AsnGlnThrSerLysLysCysGluLysValCysArgHisS    179

 221 ATGGGGAAGAATGTGCGGAGGATTTACGAAGTGCTTGATGCCTCGGGAGAG   270
     ::::::||||  |||||||    ::    :||||||
 179 erAlaSerHisThrLysGluAspLysIleGln.......ThrGlyGlu    192

 271 AAAATGAATGCTCCCACAGGACCAAGGCAGCTCGGCTTTGTCGGGAGT..   320
     ||| ::|  ||||||        |||||||:||:    ::::::
 193 LysArgLys...SerHisCysArgThrProSerLysProGluLysAlaPr   208

 321 TGGAGGCATTCGACTTCCTAACGAAAA...CTAAAGTGTGATATCTGTG    367
     ||   ::|||||||        ::||||||::      |||
 208 oGly..........SerGlyLysProTyrGluCysAsnHisCysG       220

 368 GGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCCAC   417
     ||  :::    ||:::    :::||||  ::::::|||
 220 lyLysValLeuSerHisLysGlnGlyLeuLeuAspHisGlnArgThrHis   236

 418 ACTGGAAACGGCCCTTCCAGTGCAATCAGTGCGGGCCTCATTCACCCA    467
     ||||||||::|||::::::||||||||||:|||||::|||:|
 237 ThrGlyLysProTyrGluCysAsnGluCysGlyIleAlaPheSerGl     253

 468 GAAGGGCCAACCTGCTCCGGCAACATCAAGCTGCAATTCCGGGGAGAAGCCCT   517
     |||:::||||::|:||:||  |||    ||||||||||||||||:
 253 nLysSerHisLeuValHisGlnArgThrHisArgThrGlyGluLysProT   270
```

```
 518 TCAAATGCCACCTCTGCAACTACGCCTCGCCGCCGGAGGGACGCCCTCACT   567
     |||    :||:|||          ::::::          ||||||||
 270 yrGluCysGluGlnCysGlyLysAlaHisGlyHisLysLysHisAlaLeuThr  286

 568 GGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGTGGATATTG   617
     ||||||||  |||||| ::|||::|            |||||:::::::
 287 AspHisLeuArgIleHisThrGlyLysProTyrLysCysAsnGluCy     303

 618 TGGGCCGAAGCTATAAAACACGCGAACGTCTTTAGAGGAACATAAAGACGCT   667
     ||||:::::::::    ::::::|||  :::|||      |||
 303 sGlyLysThrProArgHisSerSerAsnLeuMetGlnHisLeu...ArgS  319

 668 GCCACAACTACTTGGAAGCATGGCCGTTCCGGCACACTGTACCCAGTC    717
     |||||    |||      :::    ::          :||:|||||||
 319 erHis.ThrGlyGluLysProTyrGluCysLys..GluCysGlyLysSe  334

 718 ATTAAAGAAGAAACTAAGCACAGTGAAATGGCACGAAGAGACCTGTGCAAGAT  767
     ||:::      ::: |||||:::      :::::::|||
 334 rPheArgTyrAsnSerSerLeuThrGluHisValArgThrHisThr.Gly   350

 768 AGGATC...AGAGAGATCTTCCGTGCTGACAGACTAGCAAGTAATGTCG    814
     :::|||      ||      ||||||::|||:|||||||
 351 GluIleProTyrGluCysAsnGlyCysGlyLysGlyLysAlaPheLys...  364

 815 CCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGGGCCTG   864
     |||      |||      :||          ||:||
 365 ..............TyrGlySerSerLeuThr.................    370

 865 TCCGACACGCCCTACGAGTGCCACGTACGACGAGGAGAACGAAATGAT     914
     ||||||||            :::::::      |||||||::|||::::::
 371 ..........LysHisMetArgIleHisThrGlyGluLysProPhe      382

 915 GAAGTCCCACGTGATGGACCAAGC..........................  938
     |||          |||  :::
 383 GluCysAsnGluCysGlyLysThrPheSerLysLysSerHisLeuValI1  399

 939 .CATCAACAACGCCATCAACTACCTGGGGCCGAGTCCCTGCCGCCGCTG   987
     :|||      ||                    ThrGlyGluLysProTyrL
 399 eHisGlnArgThrHis.................ThrGlyGluLysProTyrL 411

 988 GTGCAGACCCCCGGGGGGTTCCGAGGTGGTCCGGTCATCAGCCCGAT    1037
     ::|||      |||:          :: ||||| ||
 411 ysCysAspGluCysGlyLys.......AlaPheGlyHis.SerSerSe   424

1038 GTACCAGCTGCACAGGCGCTCGGAGGGCACCCCGACCTCCAACCACTCCG  1087
                       |||:::                    HisThrG
 424 ....................rLeuThrTyrHisMetArgThr.......... 434

1088 CCCAGGACAGCGCCGTGAGTAGGTACCTGCTGCTCTCCAAGGCCAAGTTG  1137
     :::::
 434 lyAsp..............................................  435

1138 GTGCCCTCCGGAGGCGGAGGGCTCCCGGAGCAACAGCTGCCAAGACTCCAC  1187
     |||:|||                    |||:::|||
 436 CysProPheGlu..............CysAsnGlnCysGlyLysGlyLysAlaPh 447

1188 GGACACCGAGAGCAACAACGAGGAGCAGCAGCCGTCTTATCTACCTGA    1237
     |||||      :::|||                    |||
 447 e.............................LysGlnIleGluGly......LeuT 454

1238 CCAACCACATCGCCCGACGCGCCGCCAACGCCGTGTCGCTCAAGGAGGAGCAC  1287
     ||:::|||            |||
 454 hrGlnHis................GlnArgValHisThrGlyGlu......   463

1288 CGGCCCTACGACCTCGCGCGCGCCTCCGAGAACTCCAGGACGCGGCT    1337
     :::  |||||||::        |||||        ||||
 464 LysProTyrGluCysValGluCysGlyLysGlyLysAlaPheSerGlnLysSerHi 480

1338 CCGCCTGGTCAGCACC.......AGCCGGGAGCAGATGAAGGTGTACA   1378
     ::|||||||
 480 sLeuIleValHisGlnArgProHisThrGlyGlu.....AsnProPheG  495

1379 AGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACCACCATC  1428
```

seq_name: pir2:S51037

seq_documentation_block:
zinc-finger protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 01-Dec-2000
C;Accession: S51037, S05565
R;Schaefer, U.; Rausch, O.; Bouwmeester, T.; Pieler, T.
Eur. J. Biochem. 226, 567-576, 1994
A;Title: Sequence-specific recognition of a repetitive DNA element by a C(2)H(2) zinc-fi
A;Reference number: S51037; MUID:95094815
A;Accession: S51037
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-675 <SCH>
R;Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Knoech
J. Mol. Biol. 208, 639-659, 1989
A;Title: Second-order repeats in Xenopus laevis finger proteins.
A;Reference number: S05632; MUID:90040698
A;Accession: S06565
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 85-448,'K',450-613 <NIE>
A;Experimental source: clone XlcGF20-1
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C;Keywords: DNA binding; zinc finger

alignment_scores:
    Quality: 321.00   Length: 616
    Ratio: 1.092   Gaps: 21
Percent Similarity: 47.727   Percent Identity: 20.617

alignment_block:
US-08-711-417C-165 x S51037

Align seg 1/1 to: S51037   from: 1   to: 675

```
 82 GATGAGCCCATGCCGATCCCGAGGACCTCTCCACCACCTCG........ 123
    ::|||||  :::||||||||:::|||
  6 GluGluProTyrGluPheProGluAsnThrPheGlyThrAsnGluLeuLe 22

124 .GGAGGACAGCAAAGTCCAAGATGACAGAGTCGTGGCCAGTAATGTTA 172
    :::||||||  ::: :|:::|  ||:::||||:|
 22 uProAsnTyrGlnGlnAsnCysThrAspGlyAlaAlaIleSerAspThrL 39

173 AAGTAGAG....        ..ACTCAGAGTGATGAAGAG 198
    ::|  |||||||||      ||||::|:::::
 39 ysSerAspGluTyrLeuGluValGluIleThrAspAlaHisGluGlu 55

199 AATGGGCGTGCCCTGTGAAATGAATGGGGAAGAATGTCGCGAGGATTACG 248
    ::::::    |||||||||:::::| ||
 56 SerAsnThrAspLysProPheThrCysGlyCysGlyLysThrPheTh 72

249 AATGCTTGATGCCTCGGGAGAGAAAATGAAT..GGCTCCCACAGGGACC 295
    ::|||  ||| |||||| |||||:|
 72 r.................ArgLysProAsnTyrGluSerHis....... 80

296 AAGGCAGCTCGGCTTTGTCGGGAGTTGGAGGCATTCCACTTCCTAACGGA 345
    ::::|| ::||||  ::|||
 81 .................IleArgAlaHisLysGly 86
```

```
346 AAA.......CTAAAGTGTGATATCTGTGGGATCATTGCATCGGGCCCAA 389
    :::|||  :::|||
 87 GluLysProPheSerCysMetValCysAspLysAlaPheAlaTrpLysSe 103

390 TGTGCTCCATGGTTCACAAAAGAGCCACACTGGAGAACGGCCCTTCCAGT 439
    |||::||||  ::: |||::|||||| |||||
103 rAsnLeuLeuValHisTyrSerValHisSerGlyGluLysProPheSerC 120

440 GCAATCAGTGCGGGGCCTCCATTCACCCAGAAGGGCAACCTGCTCCGGCAC 489
    ||:::|||  ||||||||||||:|||  ::|||||
120 ysThrGluCysAspLysThrPheSerAsnLysAlaGlnLeuGluLysHis 136

490 ATCAAGTGCATTCCGGGAGAAGCCCTTCAAATGCCCAACCTCTGCAACTA 539
    :::::|||:|||||||:||||||||  ::|||: |||:
137 LeuArgValHisThrGlyGluLysProTyrSerCysGluGlnCysGlyLy 153

540 CGCCTGCCCGGAGGACGCCCTCACTGGCCACCTGGAGGACGCACTCCG 589
    :::|||||||||||  ::||||||||||:|||||
153 sSerPheAlaHisLysCysValLeuAspSerHisGlnArgThrHisThrG 170

590 TTGGTAAACCTCACAAATGTGGATATTGTGGCGAAGCTATAAACAGCGA 639
    |||||  :::|||  |||||||||||||||||||
170 lyAspLysProPheSerCysThrGluCysGlyLysLysPheSerGlnArg 186

640 ACGTCTTTAGAGGAACATAAAGAGCGC............ 666
    ::::|||: :::
187 GlyAsnLeuHisIsLysHisLeuLysThrHisLysLeuAspGlnProHisLe 203

667 .TGCCACAACTACTTGGAAAGCATGGCCTTCCGGGCACACTGTACCCAG 715
    |||  :::|||
203 uCysAlaGluCysGlyLysThrPheSerPheLysSerThrLeu....... 217

716 TCATTAAAGAAGAAACTAAGCACAGTGAAATG.......... 747
    |||||||||
218 ...LeuGluHisGlnLysIleHisSerGluIleArgProLeuSerGluPhe 233

748 ...........GCAGAAGACCTGTGCAAG...... 765
    |||||||||  |||
234 GlyLysThrPheSerAspAlaHisAlaAsnLeuLeuLysHisGlnSerThrPh 250

766 .........ATAG 769
    :::
250 eThrGluGlnLysProPheProCysThrGluCysGlyGlyIlePheS 267

770 GATCAGAGAGATCTCGTCGTCGGACAGA............... 798
    :::|||:||:::||:||
267 erAsnGluHisGlnSerThrHisGlnSerThrHisThrGluGluGln 283

799 ...........CTTAGCAAGTAATGTCGCCAA 818
    :::|||
284 LysProPheCysThrLysTrpGlyIlePheSerAsnGluHisGl 300

819 ACGTAAGAGCTCTATGCCTCAGAAAATTTCTTGGGACAAGAGGCCTGTCCG 868
    : |||::||||  ||| :||||:::|||
300 uLeuArgThrHisGlnSerThrHisThrGluGlyGlnLysSerLeuProS 317

869 ACACGCCCTACCGACAGTGCCACGTACCAGAAGGAGAACGAAATGATC... 915
    :::||| ::||||:|:||
317 erThrGlu...SerGlyGlyThrPheSerAsnGluHisGluLeuLeuThr 332

916 ......AAGTCCCACGTGATGGACCA............ 936
    ::|||
333 HisGlnSerThrHisThrGlyGluGlnLysHisLeuProCysThrGluCy 349

937 ........GCCATCAACAACGCCATCAACTACCTGGGGCGCCGAGTCCCTGC 979
    ::: :::|||||||||||:::|||
349 sGlyGlyThrPheThrAsnGluHisGluLeuLeuAlaHisGlnSerThrH 366

980 GCCCGCTGGTGCAGACGCCCCGGGCGGTTCCGAGGTGGTCCGGTCATC 1029
    |||  ::|||
366 isThrGluGlnLysProLeuProCysThrGluCysGlyLysIlePhe 382

1030 AGCCCGATGTACCAGCTG.......CACAGGCGCTCGGAGGGCACCCCGCG 1073
```

```
 495 :||| ::|||  |||
     luCysTyrGluCysGlyLysAlaPheAsnAlaLysSerGlnLeuValIle 511

1429 CACATGGGCTGCCACGGCTTCCGTGATCCTTTGAGTGCAACATGTGCGG 1478

 512 :||:: |||  ::: |||::|||||  ||||
     HisGlnArgSerHisThrGlyGluLysProTyrGluCysIleGluCysGl 528

1479 CTACCACCAGCCAGGACCGGTACGAGTTCTCGTCCCACATA 1518

 528 : |||||||  ||||  :|: ||||||||
     yLysAlaPheLysGlnAsnAlaSerLeuThrLysHisMet 541
```

Percent Similarity: 47.455   Percent Identity: 23.636

alignment_block:
US-08-711-417C-165 x I48689

Align seg 1/1 to: I48689 from: 1 to: 636

```
383 SerAspGluHisGluLeuThrHisGlnSerThrHisThrSerProSe 399
    |||       :::||||       |||:::   ::::||       ::::|||
 40 GGGAAGGAAAGCCCCCTGTAAGCGATACTCCAGATGAGGCGATGAGCC 89

1074 CTCCAACCACTCGGCCCCAGGACAGCGCCCTGGAGTACCTGCTGCTGCTCT 1123
     ::::::::       ::::|||                                    ||
158 GlyGluAspSerSerLeuSerThrLeuValProGlnLeuAspIleSe 174

399 rThrGluPheGlyValGlnThrThr.................... 407
    :::|||
 90 CATGCCGATCCCCGAGGACCTCTCCACCACCTCGGAGGACAGCAAAGCT 139

1124 CCAAGGCCAAGTTGGTGCCTCCGAGCGCGAGGCGGTCCCGAGCAAC... 1170
     :::::::::|||    |||||||||    |||
174 rSerSerIleArgProSerAspCysLysThrPheGlyAsnAsnLeuGluH 191

408 ..........GluAspAsnHisGlnSerProSerLysAsp 417
           ::|    :::
140 CCAAGAGTGACAGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGT 189

1171 ..........AGCTGCCAAGACTCCACGGACACCGAGAG 1199
           ||||||:::||
191 isAsnSerGluLeuValThrGlnSerAsnIleLeuAlaLysLysPro 207

418 HisThrGlyGluLysProPheSerCysSerGluCysGlyLysSerPhePh 434
    |||:::||    :::::
190 GATGAAGAAATGGGCGTGCCTGTGAAATGAATGGGAAGAATGTGCGGA 239

1200 CAACAAGAGGAGCAGCCAGCGGTCTATCTACCTGACCACC....... 1242
     :::::    |||    |||:::                |||
208 TyrLys..........Cys..............AspLysCysArgLy 215

434 eTyrLysSerValLeuLysAspHisLeuValHisThrGlyGluLysP 451
    :::::|||
240 GGATTTA.......CGAATGCTTGATGCCTCGGGAGAGAAATG...... 276

1243 .......CACATCGCCCGA..........CGCCGCCAACGCGTGCGCTC 1275
       |||         |||||:::||
215 sSerPheIleHisArgSerSerLeuAsnLysHisGluLysIleHisLysG 232

451 roTyrHisCysIleGluCysGlyArgSerTyrThrHisGlnSerSerLeu 467
    ::::::||||    ::: ||| |||
277 ............AATGGCTCCCACAGGACCAAGGCAGCTCGGCTTTG 312

1276 AAGGAGGAGCACCGCGCCTACGACCTGCTGCGCGCCGCCTCCGAGAACTC 1325
     :::::||||:::|||    ::::|| |||
232 lyAspProTyrSerAsnGlyThr.......AspGlnGly......AlaGln 244

468 LysSerHisGlnArgThrHisThrGlyValTyrValAlaPheSerCysAsnLe 484
    |||    |||:::|||        ||||||
313 TCGGGAGTTGGAGCATTCGACTTCCTAACGGAAAACTAAAGTGTGATAT 362

484 uCysAspLysLeuSerIleIleSerLysLeuArgLeuHisTyrArgValH 501
    |||    :::||    |||     |||
245 SerGly................ArgLysHisHisGluCysAlaAs 254

1351 ...................................ACC 1353
313 CTGTGGGATCATTGCATCGGGCCCAATGTGCTCATGTTCACAAAAGAA 412
     |||||    :::    ||| ||||:::|||
254 pCysGlyLysThrPheLysTrpArgThrGlnLeuThrGluHisGlnArgI 271

1354 AGCGGGGAGCAGATG........................AAGGT 1373
     :::    ||||||                                ||
363 GCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCATTC 462

501 isSerGlyGluLysProTyrProCysThrGluCysGlyLysThrPheThr 517
    |||
413 leHisThrGlyGluLysProPheGluCysAsnValCysGlyLysAlaPhe 287

518 LysLysGluGlnLeuGluSerHisTyrLysValHisThrGlyLysLysPr 534
    |||||||
463 ACCCAGAAGGGCAACCTGTCCCGGCAACCAAGCTGCATTCCGGGGAGAA 512

1374 GTACAAGTCGAACACTGCCGGTGCTCTTCCTGGATCACGTCATGTACA 1423
     |||    :::|||    |||||
271 ArgHisSerSerLeuGlyGlnHisGluAsnAlaHisThrGlyGluLy 304

534 oTyrProCysGlnGlnCysGlyLysProPheGluCysAsnValCysLeuL 551
    :::::    ::::|||
513 GCCCTCAAATGCCACCTGCGACTACGCCTCCCGCGGGGACGCCC 562

1424 CCATCCACATGGCTGCCACCGGCTTCCGTGATCCTTTGAGTGCAACATG 1473
     ::::|||        ||| |||||:::|||:::||
304 sProTyrGlnCysSerLeuCysGlyLysAlaPheGlnArgSerSerL 321

551 ysLeuHisLeuArgThrHisThrGlyAspLysProPheSerCysThrGlu 567
    :::::|||    ||||||        |||||||||:::|||
563 TCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATTGGA 612

1474 TGCGGCTACCACAGCCAGGACCGGTACGAGTTCTCGTCGCACATAACG 1521
     |||||    :::    :::::::|||||||||:::||
321 euValGlnHisGlnArgIleHisThrGlyGluLysProTyrArgCysAsn 337

568 CysGlyLysThrPheThrArgLysProAsnTyrGluSerHisLeuThr 583
    ||||||
613 TATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGGAGGAACATAAGA 662

338 LeuCysGlyLysArgSerPheArgHisSerPheThrLeuThrHisGlnHis........ 352
    ||||||||||    ||||:::|||
663 GCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCCACCGTACC 712

352 GlnVal................ 352
713 CAGTCATTAAAGAAGAAAACTAAGCACAGTGAAATGGCAGAAGACCTGC 762

353 .......GluValThrHisSerGlyGluLysProPheGlnCys 364
    |||||        :::::
763 AAGATAGGATCAGAGAGATCTCTGCTGGACAAGCTAGCAAGTAATGT 812

365 Lys.........GluCysGlyLysAlaPheSerArgCysSerSerLeuVa 378
    |||        :::::::::|
813 CGGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAG.... 858
```

seq_name: pir2:I48689

alignment_scores:
Quality: 311.00   Length: 550
Ratio: 1.192   Gaps: 22

```
alignment_scores:
       Quality:   309.00            Length:     479
         Ratio:     1.251              Gaps:        24
Percent Similarity:  51.566     Percent Identity:  24.843

alignment_block:
US-08-711-417C-165 x S10245

Align seg 1/1  to: S10245   from: 1  to: 411

130 CAGCAAAGCTCCAAGAGTGACGAGTCGTGGCCAGTAATGTTAAAGTAGA 179
    :::::::::::::::::::::     ::|||||
  6 GluArgAsnAlaThrSerTyrProThrIleSerCysLeuIleLysLysValTh 22

180 GACTCAGATGATGAGAGAGAATGGGCGTGCCTGTGAAATGAATGGGGAAG 229
    :::    |||                :::::::
 22 rLysThrLysProAlaLysAsnValArgLysTyrAlaArgHisSerAlaS 39

230 AATGTGCGGAGGATTTACGAATGCTTGATGCCTCGGGAGAGAAAATGAAT 279
    :::       :::||||||          ::
 39 erHisThrLysGluAspLysIle......GlnThrGlyGluLysArgLys 53

280 GGCTCCCACAGGACCAAGGCAGCTCGGCTTTGTCGGGAGTTGGAGGCAT 329
    ||||||         :::          ::   ::
 54 ...SerHisCysArgThrProSerLysProGluLysAlaProGly...... 67

330 TCGACTTCCTAACGGAAAA...CTAAAGTGTGATATCTGTGGATCATTT 376
    :::||||||             ||||||    :::
 68 .........SerGlyLysProTyrGluCysAsnHisCysGlyLysValL 81

377 GCATCGGCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACTGGAGAA 426
    :::::::||||||::::::||||||::::::||||||:::
 81 euSerHisLysGlnGlyLeuLeuAspHisGlnArgThrHisThrGlyGlu 97

427 CGGGCCCTTCCAGTGCAATCAGTGGGGGCCTCATTCACCCAGAAGGGCAA 476
    :::||||||:::||||||  :::|||||||:::::::|||||
 98 LysProTyrGluCysAsnGluCysGlyIleAlaPheSerGlnLysSerHi 114

477 CCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAAGCCCTTCAAATGCC 526
    :||:::|||  :::|||:::||||||:::::::||||||||||
114 sLeuValLeuHisGlnArgThrHisThrGlyGluLysProThrGluCysG 131

527 ACCTTGCAACTACGGCCTGCCGCCGGAGGACGCCCTCACTGGCCACCTG 576
    :::    |||  ::|||       ||  ||||||
131 luGlnCysGlyLysAlaHisGlyHisLysHisAlaLeuThrAspHisLeu 147

577 AGGACCGCACTCCGTTGGTAAACCTCACAAATGTGATAATTGTGGCCGAAG 626
    |||   |||     |||||||          :::
148 ArgIleHisThrGlyGluLysProTyrLysCysAsnGluCysGlyLysTh 164

627 CTATAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCGCTGCCACAACT 676
    :::::::||||||  ::::  |||||     |||
164 rPheArgHisSerSerAsnLeuMetGlnHisLeu...ArgSerHis.Thr 179

677 ACTTGGAAAGCATGGGCCTTCCGGGCACACTGTACCCAGTCATTAAAGAA 726
    :::      :::         ||||||
180 GlyGluLysProTyrGluCysLys...GluCysGlyLysSerPheArgTy 195

727 GAAACTAAGCACAGTGAAATGCAGAAGACCTGTCAAGATAGGATC.... 773
    :::  |||   :::::||||   :::::||||
195 rAsnSerLeuThrGluHisValArgThrHisThr.GlyGluIlePro 211
```

```
378 lGlnHisGluArgThr..............HisThrGlyGluLysProP 390
859 ...................GGCCTGTCCGACGCCCTAC 879
                       |||  |||   ||  ::
390 heGluCysSerIleCysGlyLysArgAlaPheGlyGlnSerProSerLeuTyr 406

880 GACAGTGCCACGTACGAGAGGAGGAACGAAATGATGAAGTCCACGTGAT 929
    ::|||::::::    ::||||:::
407 LysHisMetArgIleHisLysLysArgSerLysProTyrGlnSerAsn...... 421

930 GGACCAAGCCATCAACAACGCCATCAACTACCTGGGGCCCAGTCCCTGC 979
                                        ||||||
422 ..............AsnPhe............SerLeuA 426

980 GCCCGCTGGTGCAGACGCCCC...CCGGCGGTTCCGAGGTGGTCCCGGTC 1026
    :::  ||||   ::|||||| ||  |||:::   :::::
426 laPheValProAsnThrProLeuProGlnGlyGluGlyLeuLeuThrGlu 442

1027 ATCAGCCCGATG.............................. 1038
     ::::::
443 ValLysSerTyrHisCysAsnAspCysGlyLysAspPheGlyHisIleTh 459

1039 ....TACCAGCTGCACAGGCGTCGGAGGGCACCCCGCGCTCCAACCACT 1084
     :::  |||||:  ||  ::|||
459 rAspPheSerGluHisGlnArgLeu..............HisA 469

1085 CGGCCCAGGACAGCGCCGCTGGAGTACCTGCTCTGCTCCAAGGCCAAG 1134
     ::::::|||::: :: |||||||||:::
469 laGlyGluAsnSerTyrGlySerGluGlnThrLeuGluGlyLysGlnGlnSer 485

1135 TTGGTGCCCTCGGAGGCCGAGGCGTCCCGAGCAACAGCTGC......... 1176
     |||     |||   ||
486 LeuSerHisArgProArgGluLysProTyrGlnCysAsnValCysGlyLysAl 502

1177 ....CAAGACTCACGGACACCGAGACAACGAGGAGCAGCGCGACGCG 1222
     :::  |||||||   :::   :: ::::
502 aPheLysArgSerThrSerPheIleGluHisHisArgIleHisThrGlyG 519

1223 GTCCTATCTACCTGACCAACCACACATCGCCCGACGCGCCAACGCGTCG 1272
     |||   :::|||| ::|||||
519 luLysProTyrGluCysAsnGluCysGlyGluAlaPheSerArgLeuSer 535

1273 CTCAAGAGGACGAC.............CCGGCCTACGACCT 1301
     ::::|||  :::   |||
536 SerLeuThrGlnHisGluArgThrHisThrGlyLysProTyrGluCy 552

1302 GCTG...........CGGCGCGCCTCCAGAGAACTCGCAGGACGGCTCCGCGG 1342
     :::  ||||||  |||:::   ::
552 sIleAspCysGlyLysLysAlaPheSerGlnSerSerLeuIleGlnHisG 569

1343 TGGTCAGCACCAGCGGGAGCAGAGATGAAGGTGTACAAGTGCAACACTGC 1392
     :::  :::||||| |||  :::||||:::||||
569 luArgThrHisThrGlyGlu......LysProTyrGluCysAsnGluCys 583

1393 CGGGTGCTCTTCCTGGATCACGTCATGTACACACCCACATGGGCTGCCA 1442
     ||         ||
584 GlyArgAlaPheArgLysLysThrAsnLeuHisAspHisGlnArgThrHi 600

1443 CGGCTTCCGTGAATCCTTTGAGTGCAACATGTCGGCTACCACCAGCCAGG 1492
     :::  |||::  ||||||  |||
600 sThrGlyGluLysProTyrAlaCysLysLysGluCysGlyLysArgAsn...... 614

1493 ACCGGTACGAGTTCTCGCGACATACACGCGAGGGGAGCACCCGCTTCCAC 1542
     :::::  ||||||  :: :||||   |||
615 ....PheSerArgSerSerAlaLeuThrLysArg...HisHisArgValHis 628
```

seq_name: pir2:S10245

R;Lovering, R.; Trowsdale, J.
Nucleic Acids Res. 19, 2921-2928, 1991
A;Title: A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell
A;Reference number: S26823; MUID:91279444
A;Accession: S26823
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-803 <LOV>
A;Cross-references: EMBL:X59244; NID:g38031; PIDN:CAA41932.1; PID:g38032
R;Thiesen, H.J
New Biol. 2, 363-374, 1990
A;Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A;Reference number? I37949; MUID:91145339
A;Accession: I37967
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 476-531 <THI>
A;Cross-references: EMBL:X52358; NID:g34160; PIDN:CAA36584.1; PID:g930090
C;Genetics:
A;Gene: GDB:ZNF43; HTF6
A;Cross-references: GDB:128653
A;Map position: 19p13.1-19p12
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C;Keywords: DNA binding; zinc finger

alignment_scores:
    Quality: 308.00   Length: 405
    Ratio: 1.439   Gaps: 12
Percent Similarity: 52.840   Percent Identity: 23.704

alignment_block:
US-08-711-417c-165 x S26823

Align seg 1/1   to: S26823   from: 1   to: 803

352 AAGTGTGATATCTGTGGGATCATTTGCATCGGCCCAATGTGCTCATGGT 401
       ||||:::: |||||||
309 LysCysGluGlyCysGlyLysAlaAsnTrpProSerThrLeuThrLy 325

402 TCACAAAGAAGCCACACTGCAGAGAACGGCCCTTCCAGTGCAATCAGTGCG 451
    ||||||||| ||||||||||||||::|||::||:::||:::::|||
325 sHisLysArgIleHisThrGlyGluLysProTyrThrCysGluGluCysG 342

452 GGGCCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGTCGCAT 501
    || :||||::||| ::||||| ||| :|||||||:::|||
342 lyLysAlaPheAsnGlnPheSerAsnLeuThrThrLysArgIleHis 358

502 TCCGGGGAGAAGCCCTCAAATGCCACCTCTGCAACTACGGCTGCCGCCG 551
    ::::||||| ::||||| ||| :::||| ||| :::||::|||
359 ThrAlaGluLysPheTyrLysCysThrGluCysGlyGluAlaPheSerAr 375

552 GAGGACGGCCCCTCCACTGCGCCATGCCGAGGACGCCACTCCGTAAACCTC 601
    | ::: ||||| ||| ||||||| ||| |||||||:::|||
375 gSerSerAsnLeuThrLysHisLysIleHisThrGlyLysProT 392

602 ACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAAGGTCTTTAGAG 651
    ::||||| |||||:::|||| ::::|||||:::|||
392 yrLysCysGluGlyCysGlyLysAlaPheLysTrpSerSerLysLeuThr 408

652 GAACATAA 677
    ||||||||
409 GluHisLysLeuThrHisThrGlyGluLysProTyrLysCysGluGluCy 425

678 CTTGGAAAGCATGGGCCCTTCCGGCACACTGTACCCAGTCATTAAGAAG 727
         ::::::::: |||::::|||||
425 sGlyLysAlaPheAsnTrpProSerThrLeu 435

728 AAACTAAGCACAGTGAAATGGCAGAGAC............CTGTGCAAG 765
    |||||||||::::::: ||||||
436 ..ThrLysHisAsnArgIleHisThrGlyGluLysProTyrLysCysGlu 451

766 ATAGGATCAGAGAGATCTCGTGCTGACAGACTAGCAAGTAATGTCGC 815

774 AGAGAGATCTCGTGCTGGACAGACTAGCAAGTAATGTCGCCAAACGTA 823
    ||| :::::::||::|||| |||
212 TyrGluCysAsnGluCysGlyLysAlaPheLys 222

824 AGAGCTCTATGCCTCAGAAATTTCTTGGGACAAGGGCCTTCCGACACG 873
        |||::|||:::::::
223 ......TyrGlySerSerLeuThr 228

874 CCCTACGACAGTGCCACGTACGAGAAGGAGAACGAAATGATGAAGTCCCA 923
    ::::: :::::::::| |||||::: |||
229 ...LysHisMetArgIleHisThrGlyGluLysProPheGluCysAsn 243

924 CCTGATGGACCAAGC...............CATCAACA 946
    || ||| |||||||:
244 GluCysGlyLysThrPheSerLysLysSerHisLeuValIleHisGlnAr 260

947 ACGCCATCAACTACCTGGGGGCCGAGTCCCTGCCCCGCTGTGCGAGACG 996
    : ||| ::|||
260 gThrHis.............ThrLysGluLysProTyrLysCysSerAspG 272

997 CCCCCGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCACGCT 1046
    |||:: ::: ||||| |||
272 luCysGlyLys.......AlaPheGlyHis.SerSerSerLeuThrTy 285

1047 GCACAGGCGCTCGGAGGGCACCCCGCTCAACCACTCGGCCCAGGACA 1096
       ||| |||:::
285 rHisMetArgThr.............HisThrGlyLysAsp 293

1097 GGGCCGTGGAGTACCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCG 1146
          ||:::||| |||
294 ....CysProPhe 296

1147 GAGCGCGAGGGGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGA 1196
     |||::::||| :::
297 Glu..........CysAsnGlnCysGlyLysAlaPhe.......... 305

1197 GAGCAACAACGAGGAGGCAGCCGGTCGTTATCTACCGACCAACCACA 1246
      ::||| :::||| :::|||
306 ....LysGlnIleGluGly.......LeuThrGlnHis. 314

1247 TCGCCGACGGCGCGCAACCGTGTCGCTCAAGGAGGAGCACCGCGCCTAC 1296
     ||||||| ||| |||
315 ....GlnArgValHisThrGlyGlu......LysProTyr 324

1297 GACCTGCTGCGCGCGCCTCCAGAACTCGAGGACGGCGCCTCCGCGTGGT 1346
     | :||||| |||::|
325 GluCysValGlyCysGlyLysAlaPheGlnLysSerHisLysLeuIleHis 341

1347 C..........AGCACCAGCGGGGAGCAGGAGCAGATGAGATGGGTACAAGTGCGAAC 1387
     ||: :||||| ||| :::|||
341 lHisGlnArgThrHisGlyGlu.....LysProPheGluCysTyrArg 356

1388 ACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATGGGC 1437
     ::||| ||| ::|||||:::
356 luCysGlyLysAlaPheAsnAlaLysSerGlnLeuValIleHisGlnArg 372

1438 TGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAG 1487
     ::|||||||| ||||||
373 SerHisThrGlyGluLysProTyrGluCysIleGluCysGlyLysAlaVa 389

1488 CCAGGACGCGTACGAGAGTTCTCGTCGCACATA 1518
     ::::::::: |||:::::|||
389 lLysGlnAsnAlaSerLeuThrLysHisMet 399

seq_name: pir2:S26823

A;Accession: G02075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-595 <PON>
A;Cross-references: EMBL:U35376; NID:g1017721; PIDN:AAA79179.1; PID:g1017722
C;Genetics:
A;Gene: GDB:ZNF85
A;Cross-references: GDB:132279
A;Map position: 19p12-19p12
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

alignment_scores:
        Quality: 307.00   Length: 403
        Ratio: 1.558   Gaps: 13
Percent Similarity: 48.883   Percent Identity: 24.069

alignment_block:
US-08-711-417C-165 x G02075

Align seg 1/1   to: G02075   from: 1   to: 595

```
352 AAGTGTGATATCTGTGGGATCATTGCATCGGCCCAATGTGCTCATGGT 401
    |||||||:        |:  |   ||: :::|||::||||:
203 LysCysGluGluCysGlyLysGlyLysAlaPheAsnTrpSerThrLeuThrLy 219

402 TCACAAAGAAGCCACACTGGAGAAGCGCCTTCCAGTGCAATCACGTGCG 451
    |||||||| |||||||||:|||||||||::|||:::::::|||
219 sHisLysArgIleHisThrGlyGluLysProTyrLysCysGluGluCysG 236

452 GGGCCTCATTCACCCAGAAGGCAACCTGCTCCGGCACATCAAGTCGCAT 501
    ||  :||||:|||||||||:||||||||:||  |||::::|||
236 lyLysAlaPheAsnGlnSerAsnLeuIleLeuLysHisLysLysIleHis 252

502 TCCGGGGAGAAGCCCTCAAATGCCACCTCTGCAACTACGGCTGCCGCCG 551
    :::||||||||||:    ||:||:      |||::   ||   : ::|
253 ThrGlyGluLysProTyrLysCysGluGluCysGlyLysGlyLysAlaPheAsnAr 269

552 GAGGGACGCCCCTCACTGGCCACCTGAGGACGGACGCCACCTCCGTTGTAAACCTC 601
    |:::::||||||  |||    ||::||       |||:|        ||||||
269 gPheSerThrLeuThrHisLysIleIleHisThrGlyGluLysProT 286

602 ACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAG 651
    ::|||||   ||||||||:||:  || ::||||:::|::::::|||
286 yrLysCysLysGluCysGlyLysGlyLysAlaPheAsnArgSerSerThrLeuThr 302

652 GAACATAAA...........................GAGCGCTG 668
    ::|||:                             ||:::::|
303 ThrHisArgLysIleHisThrGlyLysGluGluCysGlyLysCysGluGluCy 319

669 CCACAAC...TACTTGGAAAGCATGGGCCTTCCGGGCCACACTGTACCAG 715
    :  || |   ::|   :::|||  :::||       |||::::|||
319 sGlyLysAlaPheLysGlnSerSerAsnLeu......ThrThrHisLysI 334

716 TCATTAAAGAAGAAAACTAAGCACAGTGAAATGGCAGAAGACCTGTGCAAG 765
    ::|||         |||::  :        ||||| :::::|||
334 leIle............HisThrGlyGluLysProTyrLysCysCysLys 345

766 ATAGGATAGAGATCTCGTGCTGGACAGACTAGCAAGTAATGTCGC 815
    :::::                          ||| ||| :::::
346 .....................AlaPheAsnGlnSe 354

816 CAAACGTAAGAGCTCTATGCCTCAGAAATTCTTGGGACAAGGGCCTGT 865
    ::: :::::||||                   ||:::|||
354 rAlaHisLeuThrHisHisGluValIleHisThrGlyGluLys....... 368

866 CCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAAATGATG 915
    |||||                         :::|||||::::::|||
369 ......ProTyrLysCys................................. 372

916 AAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGG 965
    |||:::||||:                 ::|:|||::::::::|||
373 ...............GluLysCysGlyLysGlyLysAlaPheAsnHisPheSe 383
```

```
452 :::   :::::                   |||:::::::
    ValCysGlyLysAlaPheAsnGlnPheSerAsnLeuThrHisLysAr 468

816 CAAACGTAAGAGCTCTATGCCTCAGAAATTCTTGGGACAAGGGCCTGT 865
    ::: :::::  |||    ||  ::|   |||
468 gIleHisThrAlaGluLysProTyrLys...CysGluGluCysGlyLysA 484

866 CCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAAATGATG 915
    ::: :: ::::::|||  ::||   |||::::::|||
484 laPheSerArgSerAsnLeuThrLysHisHisLysLysIleHisIleGlu 500

916 AAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGG 965
    |||:::                ::|:|||:::::::::::: ::
501 LysLysProTyrLysCysGluGluCysGlyLysGlyLysAlaPheLysTrpSerSe 517

966 GGCGAGTCCTGCCCGCCCCGTGACGCGCCCCCCGGGCGGTTCCGAGG 1015
    ::: :::  :::::::|||
517 rLysLeuThrHisGluHisLysLysIleThr.................. 527

1016 TGGTCCCGGTCATCAGCCCGATGTACCAGCTGCACAGCGGCTCGGAGGGC 1065
                                              |||
528 ..........................................Gly 528

1066 ACCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGAGTACCTGCT 1115
     ::: :::                  :::::::|||
529 GluLysProTyrLysCysGluGluCysGlyLysGlyLysAlaPheAsnHisPheSe 545

1116 GCTGCTTCCAAGGCCAGTTGGTGCCCTCGGAGCGCGAGGCGTCCCCGA 1165
     ::::||:||  :::   ::  :::::|||
545 rIleLeuThrLysHisLysArgIleHisThrGlyGluLysProTyrLysC 562

1166 GCAACAGTCGCCAAGACTCCACGGACAACGAGGACGACCAGCAG 1215
     |||:::|||  |||:::  ::: ::|||
562 ysGluGluCys...GlyLysAlaPheThrGlnSerSerAsn......... 574

1216 CGCAGCGGTCTTATCTACCTGACCAACCACCATGCCCCGACGCGCGCAACG 1265
     :::::|||              ::::|||
575 ...............LeuThrThrHis.........LysLy 580

1266 CGTGTCGCTCAAGGAGGAG...CACCCGCGCCTACGGACCTGCTGCCGCCG 1312
     ::::  |||||| |||        ::||:: ::|||
580 sIleHisThrGlyGluLysPheTyrLysCysGluGluCysGlyLysGlyLysAlaP 597

1313 CCTCCGAGAACTGCCAG...CACGCGTCCCGCGTGGTCAGCACCAGCGGG 1359
     ::::::|||:||:  |  |||  :::||::|||
597 heThrGlnSerAsnLeuThrThrHisLysLysIleHisThrGlyHisGly 613

1360 GAGCAGATCGAAGGTGTACAAGTGCGAACTGCGGGTGCTCCTTCCTGGA 1409
     |||  :::::||||||||||||::|||
614 .........LysProTyrLysCysGluGluCysGlyLysGlyLysAlaPheAsnGl 627

1410 TCACGTCATGTACCACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTT 1459
     |||:|:|| :::|||  |||   :::|||
627 nPheSerThrLeuThrLysHisLysIleIleIleHisThrGlyGluLysProT 644

1460 TTGAGTGCAACATGTGCGGCTACCAGCCAGGACCGGTAGCGAGTTCTCG 1509
     :::||:||: ||||||  |||:::|||
644 yrLysCysGluGluCysGly........LysAlaPheLysTrpSer 656

1510 TCGCACATAACGCGA 1524
     ||| :::||::|:
657 SerThrLeuThrLys 661
```

seq_name: pir2:G02075

```
                              Quality:  305.50          Length:     491
                              Ratio:      1.252          Gaps:        16
            Percent Similarity:  49.695    Percent Identity:  23.218

alignment_block:
US-08-711-417C-165 x S35305

Align seg 1/1  to: S35305    from: 1  to: 1191
```

```
 966 GGCCGAGTCCTCGCCGCCGCTGGTGCAGACGCCCCGGCGGTTCCGAGG 1015
 226 GAAGAATGTGCGGAGGATTTACGAATGCTTGATGCCTCGGGAGAGAAAAT 275
      : :::  ::: :::::::::|||
 383 rHisLeuThrThrHisLysIleIleHisGlyGlu.............. 395
 577 GluGluCys.................GlyLysAlaPh 583

1016 TGGTCCGGTCATCAGCCCGATGTACCAGCTGCACAGGGCCTCGGAGGGC 1065
 276 GAATGGCTCCCAGGGACCAAGGCAGCTCGGCTTGTCGGGAGTTGGAG 325
          :::  :::: ::|||
 396 .............LysProTyrLysCysLysGluCys 403
 583 eAsn...........HisSerSerLeuSerThrHisLysI 594

1066 ACCCGCCGTCCAACCACTCGGCCCCAGGACAGCGCCGTGAGTACCTGCT 1115
 326 GCATTCGACTTCCTAACGCAGAAAACTAAAGTGTGATATCTGTGGGATCATT 375
      ::::||||||::  : :::|||
 404 GlyLysAlaPheLysHisSerSer............... 411
 594 lelleHisThrGlyGluLysSerTyrLysCysGluGluCysGlyLysAla 610

1116 GCTGCTCCAAGGCAAGTGGTGCCCTCCGGAGGCGGAGGCGTCCCCGA 1165
 376 TGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACTGGAGA 425
          :::::||| ::::|||
 412 .ThrLeuThrLysHisLysIleIleHisThrGlyGluLysProTyrLysS 428
 611 PheLeuTrpSerSerThrLeuArgArgHisLysArgIleHisThrGlyGl 627

1166 GCAACAGTTGCCAAGACTCCACGGACACCGAGAGCAACACGAGGAGCAG 1215
 426 ACGGCCCTTCCAGTGCAATCAGTTGCGGGGCCTCATTCACCCAGAAGGGCA 475
      ::::||||::: : :::|||
 428 erLysGluCysGluLysAlaPheAsnGlnSerSerLys.......... 440
 627 uLysProTyrLysCysGlyGluCysGlyLysAlaPheSerHisSerSerA 644

1216 CGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACGCGCCAACG 1265
 476 ACCTGCTCCGGCACATCAAGCTGCATTCCGGGAGAAGCCCTCAAATGC 525
      :: :: :::::|||
 441 ...........LeuThrGluHis............LysLy 446
 644 laLeuAlaLysHisLysArgIleHisThrGlyGluLysProTyrLysCys 660

1266 CGTGTCGCTCAAGGAGGAGCACCGCGCTGCTGCGCGCCCGCCT 1315
 526 CACCTCTGCAACTACGCCTGCCCGCCGGAGGGGACGCCCTCACTGGCCACCT 575
      ::::|||:::   :::|||
 446 sIleHisThrGlyGluLysProTyrGluCysGluLysCysGlyLysAlaP 463
 661 LysGluCysGlyLysAlaPheSerAsnSerSerThrLeuAlaAsnHisLy 677

1316 CCGGAGAACTGGCAGGACGCGCTCCGCGTGTCGTCAGCACCAGGGGAGCAG 1365
 576 GAGGACGCACTCCGTTGGTAAACCTCACAAATGGATATTGGCCGAA 625
      ::::|||:::::|||
 463 heAsnGlnSerSerAsnLeuThrArgHisLysSerHisThrGluGlu 479
 677 sIleThrHisThrGluGluLysProTyrLysCysLysGluCysAspLysT 694

1366 ATGAAGGTGTACAAGTGCGAACACTGCGGGGTGCTCTTCCTGGATCACGT 1415
 626 GCTATAAACAGCGAACGTCTTTAGAGAACATAAA............... 660
      :::::::||| :::|||
 480 ...LysProTyrLysCysGluGluCysGlyLyLysGlyPheLysTrpProSe 495
 694 hrPheLysArgLeuSerThrLeuHisLysHisLysIleIleHisAlaGlY 710

1416 CATGTACACCATCCACCATGGGCTGCCCACGGCTTCCGTGATCCTTTGAGT 1465
 661 .....................GAGCGCTGC................... 669
 495 rThrLeuThrHisHisLysIleIleHisThrGlyGluLysProTyrLysC 512
 711 GluLysLeuTyrLysCysGluGluCysGlyLysAlaPheAsnArgSerSe 727

1466 GCAACATGTGCGGCTACCAGCACCGGGTACGAGTTCTCGTCGCCAC 1515
 670 ...................CACAACTACTTGGAAAGCATGGGCCTTCCG...... 699
      :::::: :::||||||:::::::::::
 512 ysGluGluCysGly..........LysAlaPheAsnGlnSerSerLys 524
 727 rAsnLeuThrIleHisLysPheIleHisThrGlyGluLysProTyrLysC 744

1516 ATAACGCGA 1524
 700 ...........GGCACACTGTACCCAGTCATTAAAGAAGAAAACTAAGCAC 738
 525 LeuThrLys 527
 744 ysGluGluCysGlyLysAlaPheAsnTrpSerSerSerLeuThrLysHis 760

 739 AGTGAAATG.........GCAGAAGACCTGTGCAAGATAGGATCAGA 776
 761 LysArgIleHisThrArgGluLysProPheLysCysLysGluCysGlyLy 777

 777 GAGATCTTCGTCGTCTGGACAGACTAGCAAGTAATGTCGCCAAACGTAGA 826
      : :::: ::  :::  :::  :
 777 sAlaPheIleTrpSerSerThrLeuThrArgHisLysArgIleHisThrG 794

 827 GCTCTATGCCTCAGAAATTTCTTGGGACAAGGGCCTGTCCGACACGCCC 876
      :::::: |||  ||| :::  :::|||
 794 lyGluLysProTyrLys...CysGluGluCysGlyLysAlaPheSerArg 809

 877 TACGACAGTGCCACGTACGAGAGGAGAAGATGATGAAGTCCCACGT 926
      :::::: |||  ||||:::
 810 SerSerThrLeuThrLysHisLysThrIleHisLysHisGlyGluLysProTy 826
```

seq_name: pir2:S35305

alignment_scores:

```
927  GATGACCAAGCCATCAACAACGCCATCAACTACCTGGGGCCGAGTCCC    976
        ::      :::::::||::::::::   ::|||    |||     :::
826  rLysCysLysGlyLysAlaPheLysHisSerAlaLeuAlaL        843

977  TGCGCCCGTGGTGCAGAGCGCCCCGGGGGTTCCGAGTGGTCCCGGTC   1026
        ::    :::::|||:||            ||::
843  ysHisLysIleIleHisAlaGlyGlu............            851

1027 ATCAGCCCGATGTACCAGCTGCACAGCGCTGCGGAGGGCACCCCGGCTC  1076
        ::|||::::     |||::  |||:::::
852  ............LysLeuTyrLysCysGluGluCysGlyLysAlaPh   863

1077 CAACCACTCGGCCCAGGACAGCGCCGTGGAGTACCTGCTGCTCTCCA    1126
     |||::||::|||:::
863  eAsnGlnSerSerAsn.............      LeuThrT        871

1127 AGGCCAAGTTGGTGCCCTCGGAGCGCGAGGCGCTCCCGAGCAACAGCTGC  1176
     |||           :::::|||       ||  :::::::::|||
871  hrHisLysIleIleHisHisThrLysGluLysProSerLysGluGluCys  887

1177 CAA...........GACTCCACGGACCACCGAGAGCAACAACGAGGA    1211
        :::           ||||||| ||||| ::::|||
888  AspLysAlaPheIleTrpSerSerThrLeuThrGluHisLysArgIleHi  904

1212 GCAGCGC...................................AGCG   1222
        |||                                     |||
904  sThrArgGluLysProTyrLysCysGluGluCysGlyLysLysAlaPheSerG  921

1223 GTCTTATCTACCTGACCAACCACCATCGCCCCGACGCGCGCAACGCGTGTCG  1272
        ::|||||:::||     |||||           |||||    |||||
921  lnProSerHisLeuThrThrHis.............LysArgMetHis     932

1273 CTCAAGGAGGAG...CACCGCGCTACGACCTGCTGCGCGCCCTCCGA    1319
     |||:::  :::::::    ::  |||   |||:::  |||
933  ThrGlyGluLysProTyrLysCysGluGluCysGlyLysLysAlaPheSerGl  949

1320 GAACTCGCAGGACGGCGCCTCCGCGTGGTCAGCACCAGCGGGGAGCAGATGA  1369
     ::|||:||   :::      |||   :::|||||:::|||||||||  |
949  nSerSerThrLeuThrThrHisLysIleIleHisHisThrGlyGlu......L  964

1370 AGGTGTACAAGTGCGGAACATGGCGAACTCCTGCTCCTGGATCACGTCATG  1419
     ||  ::||||||||::||     :::|||            ::
964  ysProTyrLysCysGluGluCysGlyLysLysAlaPheArgLysSerSerThr  980

1420 TACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTGAGTGCAA    1469
     |||  |||  |||    ||| :::::|||:::|||
981  LeuThrHisGluHisLysIleIleHisThrGlyGluLysProTyrLysCysGl  997

1470 CATGTGCGCTACCACAGCCAGACCGGTACGAGTTCTCGTCCGCACATAA    1519
        |||||||       ::::||||     ||||||  ::::|
997  uGluCysGly..........LysAlaPheSerGlnSerSerThrLeuT     1010

1520 CGCCAGGGGAGCACCGCTTCCAC   1542
     |||||  |||::::|||:::|||
1010 hrArg...HisThrArgMetHis   1016
```

seq_name: pir2:JH0500

seq_documentation_block:
zinc finger protein 29 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Dec-2000
C;Accession: JH0500
R;Denny, P.; Ashworth, A.
Gene 106, 221-227, 1991
A;Title: A zinc finger protein-encoding gene expressed in the post-meiotic phase of sper
A;Reference number: JH0500; MUID:92039080
A;Accession: JH0500
A;Molecule type: mRNA
A;Residues: 1-614 <DEN>
A;Cross-references: GB:X55126; NID:g55470; PIDN:CAA38920.1; PID:g55471
A;Experimental source: testis
C;Keywords: DNA binding; zinc finger

```
F;217-244/Region: zinc finger
F;245-272/Region: zinc finger
F;273-300/Region: zinc finger
F;301-328/Region: zinc finger
F;329-356/Region: zinc finger
F;357-384/Region: zinc finger
F;385-412/Region: zinc finger
F;413-440/Region: zinc finger
F;441-468/Region: zinc finger
F;469-496/Region: zinc finger
F;497-524/Region: zinc finger
F;525-552/Region: zinc finger
F;553-580/Region: zinc finger
F;581-608/Region: zinc finger
```

alignment_scores:
```
           Quality:   305.00          Length:           649
           Ratio:       1.105         Gaps:              26
Percent Similarity:     42.527   Percent Identity:  21.263
```

alignment_block:
US-08-711-417c-165 x JH0500

Align seg 1/1 to: JH0500 from: 1 to: 614

```
  4  GATGCTGACGAGGGTCAAGACATGTCTTTCATCAGGAGGAAGAAGCCC    53
     |||::     ||::::::      :::      |||::|||||  ||||||
  4  GATGCTGACGAGGGTCAAGACATGTCTTTCATCAGGAGGAAGAAGCCC    53

 47  AspGlyProGluSerGluProPheProGlnSerAlaGlyLysGlySerPr   63

 54  CCCTGTAAGCGATACT..........................          69
     ::|||||
 63  GlnGluGluAspAlaAlaGluGlyProGlnGlyAlaLeuValAlaArgPheA   80

 69  ..........................................          69

 80  rgGluLeuCysArgArgTrpLeuArgProGluValHisThrLysGluGln    96

 69  ..........................................          69

 97  MetLeuThrValLeuProArgGluIleGlnIleGlnAlaTrpLeuGlnHisAr  113

 70  .CCAGATGAGGGCCGATGAGCCCATGCCGATCCCCGAGGACCTCTCCACCA   118
     ||||||||||   :::  ||||||||   :::   |||||||||||
113  gProGluSerGluGluAlaValAlaLeuValGluAspLeuThrGlnT      130

119  CC................................................    120
130  hrPheArgHisSerAspPheGluIleGlnSerGluAsnGlyGluAsnSer   146

120  ..........................................          120
147  AsnGluAspMetPheGluGlyValGluSerHisGlyMetPheLeuAsnIl  163

121  .TCGGGAGGACAGCAAAGCTCCAAGAGTGACAGAGTCGTGGCCAGTAATG   169
     ||||||||||::   :::::::::::||||||
163  eSerGlyGlyGluGlyGluGlyGlnSerAsp..................    173

170  TTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGCCTGTGAAATG   219
      :::::|||||   ||||||         |||   |||
174  ...............GlyAspSerAspPheGluArgAspCysGlySer     184

220  AATGGGAAGA.........TGTCGGAGGATTTACGAATGCTTGATGC     260
     :::|||          ||||||||||  :::
185  GlyGlyAlaGlnGlyHisAlaProGlyGluAspProArgValValProSe  201

261  CTCGGGAGAGAAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCTT  310
     :::|||   |||||||||::::
201  rGluGlyArgGluVal......................GlyGlnL      209

311  TGTCGGAGTTGGAGGCATTCGACTTCCTAACGACAAAAACTAAAGTGTGAT  360
     ||   |||     :::      ||  :::  ::||||
209  euIleGlyLeuGlnGlyThrTyrLeuGlyLysProTyrGluCysPro      225
```

```
1117 CTGCTCTCCAAGGCCAAGTTGGTGCCCTCCGGAGCGCGAGGCGTCCCGAG 1166
     |||||:::              :::         :::::: |||
478  rCysGlyGluSerPheSerTrpSerSerAsnLeuIleLysHisGlnArgT 495

1167 CA...ACAGCTGCCAAGACTCCAGGACACCGAGAGCAACACACGAGGAGC 1213
     ::   |||::|  ||   |||                      :||
495  hrHisThrGlyGluLysProTyrArgCys.GlyAspCysGlyLysGlyPh 511

1214 AGGCCAGCGGGTCTTATCTACCTGACCAACCACCATGCCCGACGCGGCAA 1263
     ::|||||||||    |||            |||::|:::
511  eSerGlnArgSerGlnLeu.....ValValHisGlnArgThrHisThr.  525

1264 CGCGTGTCGCTCAAGAGGAGCACCGCGCCTACCGACCTGCTGCGCGCCGC 1313
                  |||            |||       ::::
526  ........Gly.GluLysProTyrLysCysLeuMetCysGl  536

1314 CTCCGAGAACTGCCAGAGCGCCTCCGCGGTC.......      AGCACCA 1354
     ::::: |||::  |||:::||::|||::::|
536  yLysSerPheSerArgGlySerIleLeuValMetHisGlnArgAlaHisL 553

1355 GCGGGAGCAGATGAAGGTGACAAGTGCGAACACTGCGGGTGCTCTTC 1404
     |||: |||  :||:|||  ::||| |||           ::|||
553  euGlyAsp.....LysProTyrArgCysProGluCysGlyLysGlyPhe 567

1405 CTGGATCACGTCACCATCCACCATGGCTGCCAGCGCCTTCCGTGA 1454
     |||::|||  ||           :::
568  SerTrpAsnSerValLeuIleHisIleHisGlnArgIleHisThrGlyGluLy 584

1455 TCCTTTGAGTGCAACATGTGCGGCTACCACCAGGACCCGTACGAGT 1504
     |||||    |||           ::|||
584  sProTyrArgCysProGluCysGlyLysGlyPheSerAsnSerAsnP 601

1505 TCTCGTCCCATAACGCGAGGGAGCACCGCGCTTCCACATG 1545
     ||  ::|||          :::||:::
601  heIleThrHis...........GlnArgThrHisLeu 609
```

seq_name: pir2:A38073

alignment_scores:
Quality: 304.50   Length: 528
Ratio: 1.296   Gaps: 20

```
361  ATCTGTGGGATCATTTGCATCGGGCCCAATGCTCATGGTTCACAAAAG 410
     |||||            :::     |||:::||||:::
226  GlnCysGlyLysThrPheSerArgLysSerHisLeuIleThrHisGluAr 242

411  A................................................ 411
242  gThrHisThrGlyGluLysTyrTyrLysCysAspGluCysGlyLysSerP 259

412  .............................AGCCACACTGGAGAA 426
259  heSerAspGlySerAsnPheSerArgHisGlnThrThrHisThrGlyGlu 275

427  CGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCATTCACCCAGAAGGGCAA 476
     ::||||:::||:||:|||      |||||||||||||:::
276  LysProTyrLysCysArgAspCysGlyLysSerPheSerArgSerArgAlaAs 292

477  CCTGCTCCCGCACATCAAGCTGCATTCCCGGGAGAAGCCCTTCAAATGCC 526
     |||::| |||   :|||||||||||||||::|||
292  nLeuIleThrHisGlnArgIleHisThrGlyGluLysProPheGlnCysAs 309

527  ACCTCTGCAACTACGCCTGCCCGCGAGGGACGCCCTCACTGCCCACCTG 576
     |||::| :::     |||:|| |||        ::|||
309  laGluCysGlyLysSerPheSerArgSerProAsnLeuIleAlaHisGln 325

577  AGGACGCCACTCCGTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAAG 626
     |||||||:::  |||||:::|||      ||||||
326  ArgThrHisThrGlyGluLysProTyrSerCysProGluCysGlyLysSe 342

627  CTATAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCCTGCCACAACT 676
     |:::  :::|||||||:||      ::   |||   |||
342  rPheGlyAsnArgSerLeuAsnThrHis...GlnGlyIleHis..Thr 357

677  ACTTGGAAAGACATGGGCCTTCCGGCACACTGTACCCAGTCATTAAAGAA 726
     |||:::|||||||            |||:::
358  GlyGluLysProTyrAlaCysLys...GluCysGlyLysSerPheSerTy 373

727  GAAACTAAGCACAGTGAAATGCAGCAGAGACCTGTGCAAGATAGGATCAGA 776
     :::  :::     |||::|              :::
373  rAsnSerAsnLeuIleArgHisGlnArgIleHisThr.....GlyGluL 388

777  GAGATCTCTCGTGCTGGACAGACTAGCAGTAATGTCCCAAACGTAAGA 826
     :::||                      |||:::
388  ysProTyrLysCys................................. 392

827  GCTCTATGCCTCAGAAATTTCTTGGGACAACAAGGGCCTGTCCGACACGCCC 876
     |||||:::||  :::|||      ::    |||   |||
392  ................................................. 392

877  TACCACAGTGCCACGTACGAGAGGAGAGAGAAATGATGAGTCCCACGT 926
     |||:::|||||||
393  ThrGluCysGlyGlnLysPheSerGlnSerAlaLeu............ 405

406  ..........IleThrHisArgArgThr................... 411
     |||:::|||||||
877  ..........................................GCCC 1034

927  GATGGACCAAGCCATCAACAACGCCATCAACTACCTGGGGGCCGAGTCCC 976
                    |||           |||
412  ............................HisThrGlyLysProTyrGln 419

977  TGGCCCCGCTGGTGCAGACGCCCCGGCGGTTCCGAAGTGGTCCCCGGTC 1026
     |||:::|||||  ::  |||||||  :::|||
420  CysGlyGlu..CysGlyLys..........AsnPheSerArgSe 430

1027 ATCA...........................................GCCC 1034
     ||||
430  rSerAsnLeuAlaThrHisArgArgThrHisLeuValGluLysProTyrL 447

1035 GATGTACCAGCTGC....................ACAGGCGCTCGGAGGGCA 1066
     :|||  :::                           |||
447  ysCysGlyLeuCysGlyLysSerPheSerGlnSerSerSerLeuIleAla 463

1067 CCCCGCGCCTCCAACCACTCGGCCCAGGACAGCGCCGTGGGAGTACCTGCTG 1116
     |||      |||:::                :::|||
464  HisGlnGlyThrHisThrGlyGluLys....ProTyrGluCysLeuTH 478
```

Percent Similarity: 44.508   Percent Identity: 21.402

alignment_block:
US-08-711-417C-165 x A38073   .

Align seg 1/1 to: A38073 from: 1 to: 519

```
  55 CCTGTAAGCGATACTCCAGATGAGGCGGCGATGAGCCCATGCCGATCCCGA 104
     |||  :|::::     :::||||:::||:||::|::||   ||| ||
 116 ProLysThrGluIleCysGluGluGlyAlaGluGlyAspGluIleIleSerGl 132

 105 GGACCCTTCCACCACTCGGGAGGACAGCAAAGCTCCAAGAGTGACAGAG 154
     : :|:|||      ||:::::::|| |||::|  |||:|| :
 132 uArgIleGln.......................................... 135

 155 TCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGG 204
     |||:::::: |||:||   ||| |||
 136 ..........LysAlaAspProGlnGly..ProGluLeuGly 145

 205 CGTGCCTGTGAAATGAATGGG.............................. 225
     ::|||||||||
 146 GluAlaCysGluLysGlyAsnMetLeuLysArgGlnArgIleLysArgGl 162

 226 .........AAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGC 259
     ||||::::::|||||   :::||   :
 162 uLysLysAspPheArgGlnValIleValAlaAspCysHisLeuProGluS 179

 260 CCTCGGGAGAGAAAATGAATGGCTCCCAGAGGACCAGCAGCTCGGCT 309
     :: ||:::: |||   ::   :: :::||||:::::|:
 179 erPheLysGluGluGlnLysCysLysSerGlyGlyLysTyr 195

 310 TTGTCGGGAGTTGGAGCCATTCGACTTCCTAACGGAAAACTA........ 351
     |:::||||||:::||  |||   :::
 196 SerLeuAsnSerGlyAlaValLysAsnProLysThrGlnLeuGlnLnLy 212

 352 .......AAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGC 394
     ||||:::::||||||
 212 sProPheThrCysSerValCysGlyLyLysGlyPheSerGlnSerAlaAsnL 229

 395 TCATGGTTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAAT 444
     |:::||||||:::||| |||||||||:::||||||::|||:::
 229 euValValHisGlnArgIleHisThrGlyGluLysProPheGluCysHis 245

 445 CAGTGCGGGGCCTCATTCACCCAGAAGGCAACCTGCTCCGGCACATCAA 494
     ::|||||||:: :: ||| :::|| |||:|||||:::::
 246 GluCysGlyLysAlaPheIleGlnSerAlaAsnLeuValValHisGlnAr 262

 495 GCTGCATTCCGGGGAGAGCCCCTCAAATGCCACCTCTGCAACTACGCCT 544
     :::|||:|||:::||||||   ||| |||
 262 gIleHisThrGlyLysProTyrValCysSerLysSerGlyLysLysAlaP 279

 545 GCCCGCGGAGGACGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGT 594
     ::: |||||||  |||||:::
 279 heThrGlnSerSerAsnLeuThrValHisGlnLysIleHisSerLeuGlu 295

 595 AAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTC 644
     ||||:::||| ||| :::::::::
 296 LysThrPheLysCysAsnGluCysGluLysLysAlaPheSerTyrSerSerGl 312

 645 TTTAGAGGAACATAAA...............GAGCGCTGCCACAACTACT 679
     :||||| ::|||:::|:::::
 312 nLeuAlaArgHisGlnLysValHisIleThrGlyLysLysCysTyrGluCysA 329

 680 TGGAAAGCATGGCCTTCCGGGGCCACACTGTACCCA..............GTC 717
     ||| :::||||:::
 329 snGluCys..........GlyLysThrPheThrArgSerSerAsnLeu 341

 718 ATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTGTGCAAGAT 767
     ||  ::::::  ||
 342 IleValHisGlnArgIleHisThrGlyGluLysProPheAlaCys...... 356

 768 AGGATCAGAGAGATCTCCGTGCTGGACAGACTAGCAAGTAATGTCGCCCA 817
     ::::: :: ::::::: :: ||||::
```

```
 357 ...AsnAspCysGlyLysAlaPheThrGlnSerAlaAsnLeuIleValH 372

 818 AACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGGGGCTGTCC 867
     :::|||
 372 isGlnArgSerHisThrGlyGluLys.......................... 380

 868 GACACGCCCTACGACGTGCCACGTACGACGAGGAGGAACGAAATGATGAA 917
     |||||:::    |||    :::  :::::
 381 ....ProTyrGluCysLysGluCysGlyLysLysAlaPheSerCysPheSe 395

 918 GTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGGG 967
     |||::::|| ||:::::||   |||:::::::
 395 rHisLeuIleValHisGlnArgIleHisThrAla.............. 406

 968 CCGAGTCCCTGCCCGCCTGGTGCAGACGCCCCCGGGCCGGTTCCGAGGTG 1017
     :: ||| |||||||:::|||
 407 ..............................GluLysProTyrAspCysSerGluCys 415

1018 GTCCCGTCATCAGCCCGATGTACCAGCTG.....CACAGGCGCCTCCGGA 1061
     ::::||||  :: :::|||    |||:::|||
 416 GlyLysAlaPheSerGlnLeuSerSerLeuIleValHisGlnArgIle. 431

1062 GGGCACCCCGCGCTCCAACCACTCGCCCAGGACAGCGCCGTGGAGTACC 1111
     |||||:::|| :|||
 432 ...........HisSerGlyAsp..........LeuProTyrV 439

1112 TGCTGCTGCTTCTCCAAGGCCCAAGTTGGTGCCTCCGAGCGCGAGGCGTCC 1161
     ||||:||||
 439 al................................................. 439

1162 CCGAGCAACAGCTGCCAAGACTCCACGGACGCAGAGCAACAACGAGGA 1211
     |:::|||  :::|||
 440 ...CysAsnGluCysGlyLysAlaPheThrCysSerSer........... 451

1212 GCAGCGCAGCGGTCTTATCTACCTGACCAACCATCGCCCGACGCGCGC 1261
     ||:::||::|||||:::
 452 ..............TyrLeuLeuIleHis..............G 457

1262 AACGCGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGCTGCGCGCC 1311
     |||||
 457 lnArgIleHis......................................... 460

1312 GCCTCCGAGAACTCGCAGGACGCGCTCCGGCTGGTCAGCACCAGCGGGGA 1361
     ::||||||
 461 ..............................AsnGlyGl 463

1362 GCAGATGAAGGTGTACAAGTGCGAACTGCCGGGTGCTCTTCCTGGATC 1411
     |   ||| |||:::::::|||
 463 u......LysProTyrThrCysAsnGluCysGlyLysLysAlaPheGlnA 478

1412 ACGTCATGTACACCATCCACATGGCTGCCACGGCCTTCCGTGATCCTTTT 1461
     :: |||:::||::   ::: |||:::
 478 rgSerSerLeuThrValHisGlnArgThrHisThrGlyGluLysProTyr 494

1462 GAGTGCAACATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCCGTC 1511
     ||||:::|: ::::||||
 495 GluCysGluLysCysGly............AlaAlaPheIleSerAsnSe 507

1512 GCACATAACGGCGAGGGGAGCACCGCTTCCACATG 1545
     :||||:: :::||||||  ::::::
 507 rHisLeuMetArg...HisHisArgThrHisLeu 517
```

seq_name: pir2:A40751

seq_documentation_block:
finger protein MZF1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 01-Dec-2000
C;Accession: A40751
R;Hromas, R.; Collins, S.J.; Hickstein, D.; Raskind, W.; Deaven, L.L.; O'Hara, P.; Ha
J. Biol. Chem. 266, 14183-14187, 1991
A;Title: A retinoic acid-responsive human zinc finger gene, MZF-1, preferentially exp
A;Reference number: A40751; MUID:91317761

```
A:Accession: A40751
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <HRO>
C:Genetics:
A:Gene: GDB:ZNF42; M2F-1
A:Cross-references: GB:M58297; NID:g189043; PID:g189044
A:Cross-references: GDB:125898; OMIM:194550
A:Map position: 19q13.2-19q13.4
C:Superfamily: zinc finger protein zFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; transcription regulation; zinc finger

alignment_scores:
        Quality:    304.00           Length:     527
        Ratio:        1.178          Gaps:        23
Percent Similarity:  48.956   Percent Identity:  25.806

alignment_block:
US-08-711-417C-165 x A40751

Align seg 1/1  to: A40751   from: 1   to: 485
```

```
  58 GTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCCCCAGGA 107
     ::|||   |||     :::   ::    ::  ::|||   ::|||  |
  17 IleSerAlaGlyProGlySerValSerProHisLeuHisValProTrpAs 33

 108 CCTCTCCACCACC.....TCGGGAGGACAGCAAAGCTCCAAGAGTGACA 151
     |||||::   ::   ||||||      ::   :::         |
  33 pLeuGlyMetAlaGlyLeuSerGlyGlnIleGlnSerProSerArgGluG 50

 152 GAGTCGTGGCCAGTAATGTT.........AAAGTAGAGACTCAGAGT 189
     |||   |     ::|||
  50 lyGlyPheAlaHisArgValLeuLeuProSerAspLeuArgSerGluGln 66

 190 GATGAAGAGAATGGGCGTGCCTGTGAAATGAATGGGAAGAATGTGCGGA 239
     |||     :::        |||::         |||
  67 AspProThrAspGluAspProCysArgGlyValGlyPro.......... 79

 240 GGATTACGAATGCTTGATGCTCGGGAGAGAAAATGAATGGCTCCCACA 289
     :::    :::  |||::||||| ::   :::  |||||      :::|||
  80 ......AlaAlaIleThrThrArgTrpArgSerProArgGlyArgSerA 94

 290 GGGACCAAGGCAGCTCGGCTTTGTCGGAGGTTGGAGGCATTCGACTTCCT 339
     ||   :::   |||:::              |||||   |||||
  94 rgGlyArgProSerThr.............GlyGlyGlyValValArg 105

 340 AACGGAAAACTAAAGTGTGATATCGTGGGATCATTGCATCGGGCCCAA 389
     ::|||    :::||||||:|||||||   ::
 106 GlyGly......ArgCysAspValCysGlyLysValPheSerGlnArgSe 120

 390 TGTGCTCATGGTTCACAAAAGAGAGCCACACTGGAGACGCGCCTTCCAGT 439
     :    |||::::   ||||||||||||||:::||
 120 rAsnLeuLeuArgHisGlnHisLysIleHisThrGlyGluArgProPheValC 137

 440 GCAATCAGTGCGGGGACCCTCATTCACCGAAAGGCCAACCTGCTCCGGCAC 489
     ||::||||||    |||:::||||||||||||||||||
 137 ysSerGluCysGlyArgSerPheSerArgSerSerHisLeuArgHis 153

 490 ATCAAGCTCCATTCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTA 539
     ||::||||    |||:::||||| |||:::         |||:::
 154 GlnLeuThrHisThrGlyGluArgProPheValCysHisLeuCysGlyGl 170

 540 CGCCTGCCCGCCGAGGGGACGACCCTCACTGGCCACCTGAGGACGCACTCCG 589
     :::   |||||||||||||||||||||:::|||||||||||||||||
 170 nGlyPheValArgSerAlaArgLeuGluGluHisHisArgValHisThrG 187

 590 TTGGTAAACCTCACAAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGA 639
     :::||   ||||::    ||||:::   ::|||||||
 187 lyGluGlnProPheArgCysAlaGluCysGlyLysSerPheArgGlnArg 203

 640 AGGTCTTTAGAGGAACATAAAGACCGCTCCCAACAACTACTTGAAAGCAT 689
     :::::|||    :::|||     |||:::
```

```
 204 SerAsnLeuLeuGlnHis...GlnArgIleHisGlyAspProGlyPr 219

 690 GGGCCTT..........CCGGGCCACACTGTACCCAGTCATTAAAG 724
     |||                |||||||:::         |||
 219 oGlyAlaLysProProAlaProProGlyAla.....ProGluProProG 234

 725 AAGAAACTAAGCACAGTGAAATGGCAGAAGACCTGTCAAGATAGGATCA 774
     |||||          ||::::::::::::
 234 lyProPheProCysSerGluCysArgGluSerPheAlaArg........ 247

 775 GAGAGATCTTCGTGCTGGACAGACTAGCAAGTAATGTCGCCAAACGTAA 824
     |||::||||| |||    |||   :::     :::::
 248 ...ArgAlaValLeuLeuGluHisGlnAlaValHisThrGlyAspLysSe 263

 825 GAGCTCTATGCCTCAGAAATTTCTTGGGACAAGGGCCTGTCCGACACGC 874
     :: |||  ||:::   ::: ||               |||
 263 rPheGly.CysVal.GluCysGlyGluArgPheGlyArgArgSerValLe 279

 875 CCTACGACAGTGCCACCTACGAGAAGGAGAACGAAATGATGAAGTCCCAC 924
     |||::|||   ::||:::||||||||||                   :
 279 uLeuGlnHisArgArgValHisSerGlyValGluArgProPheAlaCysAlaG 296

 925 GTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGGGCCGAGTC 974
     :: || |||   ::||| ::::|||   ||                ||
 296 luCysGlyGlnSerPheArgGlnArgSerAsnLeuThrGlnHisArg... 311

 975 CCTGCGCCGCTGGTGCGAGAGCCCCGGGGTTCCGAGGTGGTCCGG 1024
     ||              ||:::||||||
 312 .................ArgIleHisThrGl 316

1025 TCATCAGCCCGA...TGTACCAGCTGCACAGGCGCTCGGAGGGCACCCC 1070
     :::::::|||             ||| :::||||
 316 yGluArgProPheAlaCysAlaLysCysGlyLysLysAlaPheArg.... 330

1071 GCGGCTCCAACCACTCGGCCCAGGACGACGCCGCCTGGGAGTACCTGCTGCTGC 1120
     |||||    :::          :::
 331 ..........GlnArgProThrLeuThrGlnHisLeuArgValHisThr 343

1121 TCTCCAAGGCCAAGTTGGTGCCCTCCGGAGCGGCGAGGCGTCCCCGAGCAAC 1170
     ||:::|||||||    ::|||||| :::|||:|
 344 GlyGluLysProPheAlaCysSerProGluCysGlyGlnArg.PheSerGlnA 360

1171 AGCTGCCAAGACTCCACGGACACCGAGAGCAACAACGAGGAGCAGCGCCAG 1220
     ||||||    :::|||   ::||||||   |||
 360 rgLeu...LysLeuThrArgHis...............Gln 367

1221 CGGTCTTTATCTACCTGACCAACCACAT................. 1247
     ||:::|||:::||:||||:::
 368 ArgThrHisThrGlyGluLysProTyrHisCysGlyGluCysGlyLysLeuGl 384

1248 .........CGCCCCAGCGCGCCAACCGGTGCTGCTCAAGGAG 1281
     |||     |||   ::|||::||
 384 yPheThrGlnValSerArgLeuThrGluHisArgIleHisThrGlyG 401

1282 GAGCACCGCCGCCTACGACCT...GCTGCGCGCGCCTCCGGAACTCGCA 1328
     |||          ||||||||::: ::|||||||   |||
 401 luArgProPheAlaCysGlyCysGlyGlnSerPheArgGlnHisSerAla 417

1329 GGACGGCGCTCCG...CGTGGTGCAGCGCCAGCGGGAGCAGATGAAGGTGT 1375
     :::  :::           |||||| ::|||||||
 418 ...AsnLeuThrGlnHisArgArgIleHis.ThrGlyGlu......ArgProF 432

1376 ACAAGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACACC 1425
     || |||    ||:::||   |||||||||| :::|||
 432 yrAlaCysProGluCysGlyLysLysAlaPheArgGlnArgProThrLeuThr 448

1426 ATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTGAGTCGAACATGTG 1475
     |||   :::|||||||   ||||| :::|||   |||
 448 GlnHisLeuArgThrHisArgArgGluLysProPheAlaCysGlnAspCy 465

1476.CGGCTACCACAGCCAGACCGGTACGAGTTCTCGTCCCACATAACCGCGAG 1525
     |||
 465 sGly. 466
```

Wed Aug 28 10:05:20 2002

us-08-711-417c-165.rpr

```
1526 GGGAGCACCGCTTCCACATGAGC 1548
     :::||||||||:::|||
467 .....ArgArgPheHisGlnSer 472
```